

## Bad peer reviewers

**A small proportion of referees are undermining the scientific process, especially in biology. Some of the problems are getting worse, partly because of changes in scientific publishing.**

**M**ost researchers agree that peer review is the least imperfect way of upholding the quality of scientific publications. But those who administer it also have to cope with, and attempt to solve, the problems that peer review gives rise to.

One such problem is misconduct. The peer-review process depends on trust, and the great majority of reviewers are trustworthy, but there are occasions when that trust has been abused by a referee. One of the worst that *Nature* has experienced, some years ago, was when a referee obstructed a paper, and used its information indirectly to obtain materials from the author to complete competing work of his own, which he then promptly published, thereby scooping the original. (The referee has not been used again.)

Like most journals, *Nature* requests that referees disqualify themselves from refereeing a paper that is directly competitive with their own work. We also request that referees either keep the paper strictly to themselves or consult, also in confidence, a close colleague whom they identify to us. We use two and often three referees, and do what we can (which we believe is quite a lot) to spot phoney arguments against a manuscript and to sidestep undue delay.

But there is only so much that *Nature* or any journal can do. Take confidentiality: as can be seen in a survey of the problems (see page 102), scientists vary in their attitudes to the principle of strict confidentiality. Some freely consult colleagues to an extent that others (including some authors) would find horrifying. And it is not unknown for a *Nature* editor to visit a lab and see a manuscript sent in confidence left casually on display on the group coffee-table.

Even more concerning is a discernible growth in cases where referees sit on manuscripts, or are subsequently shown (or strongly suspected) to have plagiarized them, or used them as a stimulus to rush their own, sometimes less-developed versions, into print in a competing journal. On the rare occasions where the offence is undeniable, *Nature* can take sanctions, for example by contacting the offender's employer.

Misconduct in peer review is very infrequent. But in the most competitive areas of biology — molecular biology, in particular — it is, in *Nature's* experience, no longer unexpected. In seeking to address this regrettable situation, it is important to examine underlying pressures.

The particular competitiveness in molecular biology appears to stem from a number of factors, not the least of which are the rampant egotism in the upper echelons of the field and the urgent need to publish if postdocs and contract researchers are to obtain grants. Furthermore, the route from basic research to commercial exploitation is a particularly short and direct one. Duplication of effort is yet another factor. The risk that years of painstaking work will be reduced to insignificance through being scooped by competitors seems significantly greater in molecular biology than in other disciplines. This is partly because of the inexorable pressure on researchers to attack problems guaranteed to yield quick and safe scientific returns.

With the best of intentions, science publishers are now inadvertently exacerbating the situation. With every encouragement from researchers, more and more journals (including *Nature* and its related titles) are moving in the direction of electronic publication ahead of print. This can be good for science, and especially so where the research has immediate positive implications for human health, for example. But everyone suffers if the practice is allowed to encourage corner-cutting in research and to further stimulate bad behaviour.

Perhaps it is time to review the concept of the scientific scoop. Many journals, including this one, are tough on authors of papers under revision whose results are scooped by others. But papers reporting similar results published weeks apart will all be found by an electronic search of the literature, and the later publications may be of higher quality by virtue of the additional time taken. Maybe there is a better balance to be struck between the recognition of quality and of priority than is currently the norm. ■

## Astronomy's real priorities

**Debates over the administration of US astronomy funding have highlighted areas for collaboration between agencies.**

**A**stronomers in the United States may feel that they have won a round after a National Academy of Sciences panel gave the thumbs down to a White House suggestion — tentative though it was — that federally funded astronomy programmes be consolidated within NASA (see page 99). There are many good reasons to keep the planning and management of ground-based telescopes within the National Science Foundation (NSF), not the least of which are its close ties to the academic research community.

And the panel's call for a new interagency coordinating body for astronomy could lead to something the White House may not have anticipated when it commissioned the academy study — an ongoing, high-profile forum for astronomers to plead for more money. That, after all, is big astronomy's biggest problem, rather than a lack of coordination between NASA and the NSF. True, communication

between the two agencies could be better. But even though they work in separate arenas (ground-based and space-based), the academy panel concluded that the NSF and NASA have done a good job of implementing the priority projects outlined in the astronomy community's 'decadal surveys', which are themselves models of consensus-building.

Certainly, it would be good for astronomers, through standing advisory groups, to help oversee how NSF implements the even larger telescope projects on the drawing-board. And NSF and NASA should develop a joint plan for explaining astronomy to the public.

But it still all comes down to money. If they want world-class results, the Office of Management and Budget and Congress will have to allocate the dollars for more of those expensive instruments that have led, in the words of the academy panel, to the current "extraordinary period of scientific progress" in astronomy and astrophysics. ■





**Meeting of minds**  
Brazil seeks fresh focus for science and technology  
p96



**Crime wave**  
US research vessel has a narrow escape from armed pirates  
p97



**The way ahead**  
Scotland plans to better exploit its scientific expertise  
p98



**Ill-starred plan**  
Science panel says no to merger in US astronomy  
p99

# French researchers take a stand against cancer gene patent

**Declan Butler and Sally Goodman, Paris**

The Curie Institute in Paris is to challenge a European patent held by the US biotechnology company Myriad Genetics. The patent covers the *BRCA1* gene, used in tests to assess a patient's predisposition to hereditary breast and ovarian cancers.

The French government is officially supporting the institute's action and intends to introduce legislation that would extend compulsory licensing to cover genetic tests.

Utah-based Myriad Genetics, which won a US monopoly on the *BRCA1* gene in 1999, obtained a European patent on it in January. As a result, Europeans will be obliged to pay \$2,400 for Myriad's screening test — alternative French tests cost less than a third of this.

Worldwide, Myriad has aggressively exercised its exclusive rights to carry out diagnostic testing on *BRCA1* and a related gene, *BRCA2*. It requires that all samples for testing be sent to the company's labs in Salt Lake City.



**Health worries:** Dominique Stoppa-Lyonnet says Myriad's test did not identify some expected mutations.

Estimates vary as to the proportion of breast cancer cases that are hereditary, but the US National Cancer Institute suggests

that *BRCA1* is responsible for about half of those cases that are genetically inherited.

Myriad's patent, EP 699754, covers all methods for diagnosing the risk of breast cancer based on comparing a sample sequence with the sequence it describes for *BRCA1*. At a press conference in Paris on 6 September, the Curie Institute said that it will challenge the patent on grounds of lack of novelty and inadequate description.

The institute argues that the patent draws on 'prior art' generated by public genome centres, and also that the sequence used in the initial patent application contained errors, limiting its usefulness. Jacques Warcoï, the Curie Institute's patent adviser, says that other European countries, including Germany, may join the challenge.

In addition, Dominique Stoppa-Lyonnet, a physician responsible for genetic testing of breast cancer at the Curie Institute, claims that the automated sequencing method used by the Myriad test to identify deletions and mutations detects only 10–20% of the expected mutations.

Stoppa-Lyonnet was co-author of a paper published in June (*J. Med. Genet.* **38**, 388–391; 2001) describing a deletion in the gene accounting for the predisposition of one US family with a history of breast and ovarian cancer. The Myriad test failed to detect any *BRCA1* predisposition in the family, the

## Venture capital concerns academics

**Rex Dalton, San Diego**

The University of California, San Francisco, (UCSF) has set up a scheme to fund projects aimed at generating commercial products such as drugs or equipment.

Established by Burrill, a San Francisco venture-capital firm, the scheme will provide grants of up to \$500,000 for high-risk research projects that might be unlikely to attract grants from public agencies.

If the projects are successful, Burrill would help to set up a company to exploit them, which it could own jointly with the UCSF, the faculty members involved, and the California Public Employees' Retirement System pension fund, which is making an initial investment of \$10 million.

But the arrangement was agreed by UCSF administrators without formal consent from the university faculty, some of whom are concerned about its implications for academic freedom.

The deal raises "huge potential conflict-

of-interest issues", says Stanton Glantz, a cardiovascular researcher and chair of the UCSF academic senate planning committee. "This could have major impacts on the whole mission of the university," he says.

"Unfortunately, the administration felt pressured to sign the agreement before adequate vetting by the faculty," says Daniel Bickle, an endocrinologist and chair of the academic senate.

Zach Hall, the UCSF's vice chancellor for research, says money from the deal will go back into research. "Part of our mission as a state public institution is to make sure discoveries in labs are converted to products for public good," he says.

The UCSF is building a biomedical research campus, Mission Bay, at which academic and privately funded researchers are expected to work in close collaboration. The university hopes that deals such as the one with Burrill will keep it in the forefront of partnerships with the private sector. ■

► authors said. They detected the new deletion — which, with 12,000 base pairs, accounts for about 15% of the gene — using a technique patented by the Pasteur Institute called combed DNA colour bar coding.

William Hockett, a spokesman for Myriad, says that the company is developing a method to detect such large deletions and that a commercial test will be available by the end of the year.

Opposition to a patent must be filed within nine months of the date the patent was granted, and Warcoin says they will wait until the 9 October deadline before filing to ensure their case is as strong as possible. Hockett says Myriad is unaware of the planned opposition, and would not comment until it had been filed.

The Curie Institute argues in particular that the requirement for samples to be sent to Myriad for testing will deny French scientists data and expertise, and will hamper development of new tests.

The opposition will also challenge the patent's breadth. As it covers all diagnostics based on the gene sequence, it allows Myriad to prevent the marketing of tests developed by the Curie Institute, and others, that use different techniques from Myriad's test.

Roger-Gérard Schwartzberg, the French research minister, said last week that the government would introduce a bill to extend the laws that govern compulsory licensing to cover diagnostics, and consider exercising this power if Myriad refuses to license its tests to be carried out in France. ■

## Helmholtz Society prepares itself for strategic reforms

Quirin Schiermeier, Munich

The Helmholtz Society, Germany's largest government-funded research organization, is expected to agree to sweeping reforms at its annual meeting in Berlin this week.

The reforms will introduce greater centralized control for the society's 16 research centres. The government proposed the changes earlier this year in an effort to increase the centres' relevance to industry and society, and to foster more collaboration and competition between the various sites.

The reform package will give the society's president and senate significantly more power to determine research goals and allocate funds, based on the result of scientific evaluations every five years.

But many Helmholtz researchers worry that this centralizing drive could undermine basic research and curtail their scientific independence.

Under the reforms, most of the organization's DM2.6-billion (\$US1.2-billion) annual budget will be earmarked to six strategic programmes. Researchers say that this will make it hard for them to get external grants from funding agencies, such as the European Commission, which expect host institutions to cover infrastructure and overhead costs.

In response to protests earlier this year,



**Walter Kröll: the centres will keep their autonomy.**

research minister Edelgard Bulmahn agreed to include "basic research" as one of the missions to be pursued by the society. "This is an important rider," says Rudi Balling, scientific director of the German Research Centre for Biotechnology in Braunschweig, "but only time will tell if it keeps our scientific freedom alive."

Walter Kröll, director of the German Space Agency, who is likely to be named in November as the Helmholtz's next president, defends the reforms. Research planning according to public demands "is not incompatible with scientific freedom", he says. "The centres will be left with enough autonomy to develop their own scientific profiles."

The first Helmholtz Society centres were set up in the 1950s to research atomic energy. Over the years, their activities have diversified and they now employ 8,000 researchers between them, and cover areas such as molecular genetics, cancer research, plasma and particle physics, space technology, and environmental, marine and polar research. ■

## Summit to put education at the heart of Brazil's future

Ricardo Bonalume, São Paulo

A fresh focus for Brazilian science is likely to emerge from a summit to be held on 18–21 September in Brasilia.

Organized by the Ministry of Science and Technology, the National Conference of Science, Technology and Innovation will attempt to determine Brazil's science and technology policy for the next 10 years.

According to a 250-page green paper written to help guide the meeting, policy should move on from the traditional focus of research investment to tackle two areas of perceived weakness: general levels of public education and industrial innovation.

"The great success of the 'Asian tigers' came largely from the role that families give to the education of their children," says physicist Cylon Gonçalves da Silva, former director of Brazil's national synchrotron at Campinas, near São Paulo. But the level of schooling for the average Brazilian has grown only slowly — from four years to six — since 1981, he points out.

Da Silva coordinated the 400 scientists

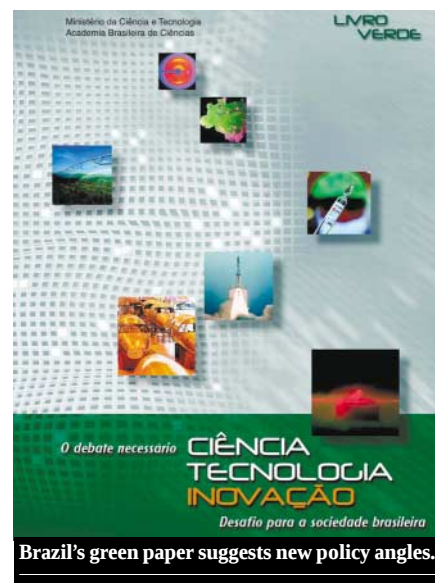
involved in writing the green paper and helped to plan the conference's programme. The summit will involve representatives of scientific societies, industry and agriculture — the first such meeting since 1985. In the intervening years, Brazil has invested heavily in training PhDs and research, and has sharply increased its share of publications in leading international journals.

But a government freeze on university hiring means that the country is struggling to employ its new graduates. Writing in *Nature* (413, 16; 2001), for example, four Brazilian researchers warn that "a whole generation of young PhDs will be lost to research".

"The idea of the conference is a very good one," says Glaci Zancan, a biochemist at the Federal University of Paraná and president of the Brazilian Society for the Advancement of Science. He says that it should focus on how science can help to address inequity in Brazilian society.

The meeting will help to produce a white paper on policy. But Brazil will elect a new president next year, and some fear that the

new administration may ignore the paper, unless the meeting reaches a strong consensus on what it should say. ■





# Pirates attack US research ship off Somalia

**Mark Schrope**

A US oceanography research ship has been attacked by pirates in the Gulf of Aden — leading its managers to rethink their plans for studying salinity flows in the region.

Unknown, heavily armed assailants pursued the *Maurice Ewing* research vessel in a small boat for a harrowing 20 minutes in waters 30 kilometres off the coast of Somalia on 31 August. No crew members were hurt and the ship was not damaged.

The *Maurice Ewing*, which is operated by the Lamont-Doherty Earth Observatory at Columbia University, New York, was in the region for a project called the Red Sea Outflow Experiment (REDSOX). The project involves researchers from the Woods Hole Oceanographic Institution in Massachusetts and the University of Miami in Florida.

"I was frightened, and in a state of disbelief," says Amy Bower of Woods Hole, chief scientist on the mission. "Although I was well aware that this was a region known for piracy, I couldn't believe that our vessel would ever come under fire."

When it first became apparent that the men approaching in a small boat were potentially dangerous, Bower says, all the scientists on the vessel gathered in its main laboratory. "From inside the ship, I don't think anyone heard any shots," she says. "There's just too much background noise — but we watched on a video monitor as crew members on the back deck ducked and we heard over the ship's radio that shots had been fired. It was becoming a much more serious situation, and everyone was ordered to go to their staterooms and lock the doors."

Fire hoses were prepared to repel boarders, Bower says, but the boat gave up the chase.

Mike Purdy, director of Lamont-Doherty, says the expedition was planned in accordance with Department of State security guidelines. These guidelines recommend that ships stay at least 19 kilometres off Somalia's coast, because of the risks of piracy and terrorism. Other research vessels have operated recently in the same area, he says.

"We analysed the risks and made the judgement that it would be a reasonable operating area. We were clearly wrong in that judgement," says Purdy. Nonetheless, he says, the ship's crew handled the situation well, executing predetermined security procedures. He says the crew thinks the attackers fired a grenade, which failed to detonate. The incident has been reported to the state department and the US Navy, but Purdy has no information on the reason for the attack.

As part of REDSOX, researchers completed another cruise in the area in March on the *Knorr*, operated by Woods Hole. Their goal is to complete the first comprehensive survey of the movement of high-salinity water from



In calmer waters: shots and possibly also a grenade were fired at the *Maurice Ewing*.

the Red Sea as it mixes with waters in the Gulf of Aden, and eventually the Indian Ocean.

Following consultation with the National Science Foundation, which funded the mission, and others, Purdy authorized the ship to continue working but to stay at least 80 kilometres off the coast. The cruise's first leg is scheduled to end this week in Djibouti.

Extensive work closer to the coast had been scheduled for the second leg, but, says Purdy, "our number one priority right now is the safety of the vessel and the folks on the vessel. We are making substantial modifications to the science plans to ensure that the vessel is never again placed in the dangerous situation in which it found itself." ■

## WHO plans study of Gulf War fallout

**Alison Abbott, Munich**

The World Health Organization (WHO) is to work with Iraqi scientists to assess claims that the incidence of certain diseases has increased in Iraq as a result of NATO's use of weapons containing depleted uranium.

A group of WHO experts visited Iraq late last month to discuss how best to design a study to investigate the claims.

Iraqi scientists say that the use of depleted uranium during the 1991 Gulf War has led to an increased occurrence of at least six types of cancer, and to changes in their characteristics. Renal disease and congenital malformations have also increased, they claim.

Abdelaziz Saleh, from the WHO regional office in Cairo, led the visiting group. "It is not possible to comment on the Iraqi data yet," he says. "But we are preparing a good study design that will allow the questions to be answered definitively."

But definitive answers could be difficult to obtain. An expert on depleted uranium at the Vienna-based International Atomic Energy Agency says the situation in Iraq is an "epidemiological nightmare".

Not only is it unclear whether Iraq has a reliable basis for collecting health data, the expert says, but there are many possible agents that could influence health statistics in postwar Iraq, including poor nutrition and other contaminants left by the war, such

as residues from burning oilfields. Moreover, he says, there is no clear account of where the uranium used in the war was deployed.

Barry Smith, of the British Geological Survey, who helped to prepare a WHO report on depleted uranium published in April, says that epidemiology is a useful tool that, together with risk assessment, can help to answer these questions. But such studies require careful investigation of contaminated environments and any potential routes of human exposure. "Both of these can be particularly difficult in post-conflict situations," he says. ■



A WHO team is in Iraq to examine health patterns in the aftermath of 1991's Gulf conflict.

LAMONT-DOHERTY EARTH OBSERVATORY

AP/US NAVY

## Relaxed restrictions blamed for rise in foot and mouth

Jim Giles, London

Epidemiologists working on the British foot-and-mouth epidemic are blaming the relaxation of security measures for a burst of small outbreaks of the disease.

The epidemic had seemed to be under control in May (see M. Woolhouse *et al.* *Nature* 411, 258–259; 2001). The number of new cases reported each day was below five for most of June, July and August, down from a peak of more than 50 in April, and government officials hoped that the disease would be extinct by the end of the year.

But in Northumberland, more than 20 cases have been recorded since late August. Before that, no new cases had been reported there since May. “The tail of the disease is longer than had been hoped,” says Mark Woolhouse, an epidemiologist at the University of Edinburgh.

Woolhouse points to the relaxation of restrictions on moving animals and measures to stop people carrying the virus between farms. The government had allowed some movement of animals on animal-welfare grounds. Also, farmers and vets may have stopped applying restrictions so rigorously, scientists say.

Researchers say there is a chance the disease surfacing in Northumberland was carried in sheep flocks, as sheep can carry the virus without showing any symptoms.

The government has now halted animal movements in affected areas and has imposed stricter security measures. But this is causing problems for farmers in upland areas, such as Northumberland. Animals that graze on high ground in summer are normally relocated at this time of year.

Epidemiologists say it is difficult to estimate how long the epidemic will last. The onset of winter may cause problems, they add, as the virus can survive for longer in lower temperatures. ■



## Scotland aims to capitalize on spirit of scientific innovation



Looking to the future: Wendy Alexander seeks a smarter approach to supporting science.

David Dickson, Glasgow

Two years after re-establishing its own parliament in Edinburgh, Scotland has unveiled a science strategy aimed at bolstering innovation and increasing public understanding of science.

As part of the plan, released last month by Wendy Alexander, the Scottish Executive's minister for enterprise and lifelong learning, an advisory committee will be set up under the auspices of the Royal Society of Edinburgh (RSE) to propose research priorities and other measures for boosting Scottish science.

The strategy also pledges to enhance the teaching of science in schools, promote public understanding of science and create “a pipeline of support to carry science from Scottish laboratories into the creation of global companies”.

Alexander, whose brief includes responsibility for science, said the country has a “proud history of innovation” and an “enviable international reputation in important new fields”, referring to areas such as biotechnology and opto-electronics. But she added that the government “must become smarter still in the support it gives to science, the use it makes of science, and the way it explains the issues”.

Her remarks reflect concerns both over Scotland's slowness to benefit from its own scientific discoveries and at the potential negative impact of public worries about breakthroughs such as the cloning of Dolly the sheep at the Roslin Institute, near Edinburgh.

Eddie Frizzell, the senior civil servant in Alexander's department, told the British Association for the Advancement of Science's

meeting in Glasgow last week that “too often scientific knowledge produced in Scotland is exploited in other parts of the world”. He added: “We are going to ask the advisory group at an early stage to tackle the question of where the priority areas should be.”

This view was endorsed by Sir William Stewart, president of the RSE and a former chief scientific adviser to the British government. “We cannot do everything and need to focus, particularly on areas that will give this country a competitive edge,” he said.

The Scottish government spends £220 million (US\$320 million) on research and development each year, and Scotland obtains another £165 million a year from the British government, mainly in grants to universities from UK research councils.

The science strategy was generally welcomed by Scottish researchers, although some complain that it lacks specifics. “This is a framework rather than a route map,” admits Geoffrey Boulton, professor of geophysics at the University of Edinburgh and a member of an RSE committee that influenced the strategy's recommendations.

But advocates of greater public involvement in science said the strategy is too conservative. “I am very pleased to see a science strategy for Scotland, but these things need to be given a higher priority,” said Donald Bruce, a former nuclear chemist who now heads the Church of Scotland's Science, Religion and Technology Project. The document “could have been written 10 years ago”, he said. “It emphasizes factors such as the need for scientists to communicate more with the public, but says very little about the need to listen to where the public is coming from.” ■



# Astronomers buoyed by rejection of merger

**Tony Reichhardt, Washington**

Astronomers have welcomed the rejection by a top level panel of a proposal to subsume into NASA the ground-based astronomy programme run by the National Science Foundation (NSF).

The White House had suggested earlier this year that NASA should take control of the programme as a way to reorganize federal funding for astronomy, but the plan was unpopular among scientists from the start. Researchers feared, among other things, that the mission-oriented space agency would not be as guided by scientific priorities as the NSF (see *Nature* **410**, 853; 2001).

After carrying out a fast-track study ordered by the White House, the panel from the US National Academy of Sciences — chaired by Norman Augustine, former chief executive of Lockheed Martin — flatly rejected the idea. Its report instead recommends establishing a high-level joint advisory committee to supervise both agencies' astronomy programmes.

"The NSF is the right institution to sponsor ground-based astronomy and astrophysics," the study concludes, adding that a simple transfer to NASA "would have a net disruptive effect on scientific work". Astronomers are "pretty happy" with this unambiguous conclusion, says Kevin Marvel, a policy specialist with the American

Astronomical Society in Washington. The society, whose members sent hundreds of e-mails to the panel during the study period, is preparing its formal response to the report.

Instead of creating a single 'super-agency' for astronomy, Augustine's panel recommends that the White House science office and its Office of Management and Budget create a new interagency planning board for astronomy and astrophysics, with membership drawn from the energy and defence departments, the Smithsonian Institution and other secondary players in astronomy, as well as from NASA and the NSF. The coordinated interagency programme could be modelled on the US Global Change Research Program, it suggests.

The NSF's astronomy division would benefit from having a permanent advisory committee like NASA's, the panel says, which would help allay concerns that the NSF is struggling to manage ground-based astronomy projects, including the \$600-million Atacama Large Millimeter Array (ALMA) radio telescope planned in Chile. "Several staff members in the Executive Branch and in Congress conveyed to the committee their perception that NSF does not manage large projects well," says the report.

Yet this may be just a perception. Augustine's group "did not find evidence" that the NSF had significantly more problems during

the construction phase of large projects than did NASA, the Department of Energy, or other similar agencies. But the NSF has communicated its strategies for these projects poorly to the White House and Congress, says the report — a complaint echoed last week in a House of Representatives science committee hearing on the NSF's management of large research facilities.

Even with more regular advice from the scientific community, and after strengthening its communication with NASA, the NSF will face an uphill battle in fulfilling astronomers' wishes for large new telescopes, as sources of funding have not been identified. "By a substantial margin, the NSF does not have the resources to keep US ground-based optical and infrared astronomy at the world level," the study panel says. ■



The NSF will still run Puerto Rico's Arecibo dish.

# NIH ponders repository for embryonic stem cells

**Laura Bonetta, Washington**

The US National Institutes of Health (NIH) is considering plans for a repository to house human embryonic stem cells for distribution to US researchers.

The proposed stem-cell bank would hold many of the 64 embryonic stem-cell lines approved for funding under the Bush administration's recent ruling (see *Nature* **412**, 665; 2001).



Tommy Thompson: concedes that fewer than half of the approved lines are ready for use.

Health secretary Tommy Thompson unveiled the plan at the 5 September hearing of the Senate's Health, Education, Labor and Pensions Committee. Thompson also conceded at the hearing that only 24 or 25 of the 64 approved cell lines are fully developed and "ready to be sent out to researchers".

The repository idea is an extension of the NIH's plan to compile a registry of embryonic stem cells, which will initially include the name of each cell line and contact information for the provider. Supporters of the concept say that a repository might ease researchers' concerns about the quality and accessibility of the cell lines that are eligible for funding.

Although the concept is supported by many scientists and scientific organizations, it is unclear whether providers of embryonic stem cells will agree to it. "Some of the entities that have derived stem cells have expressed interest in a repository, others have not yet addressed this idea, and still others have said they are not interested," says Lana Skirboll, associate director for science policy at the NIH.

Douglas Melton, a stem-cell researcher at Harvard University and an advocate of the repository, told the Senate hearing: "The federal government and the NIH are in an immeasurably stronger position than are individual investigators to obtain the human embryonic stem-cell lines from suppliers, verify their quality, and arrange for their distribution."

The idea of a human embryonic stem-cell bank is not new. In March 2000, a working group for the Royal Society recommended creating a repository in the United Kingdom, where regulations on stem-cell research are less restrictive than in the United States. In a more recent report, the society said that cell lines would be available to the scientific community on a non-commercial basis, and "funding and ethical permission on stem-cell research could be conditional on any lines being put in the bank".

The Senate committee's hearing was the first of three scheduled for this month, as senators probe the justification for the Bush administration's policy on the funding of stem-cell research. ■

## Norway seizes on century-old proposal for top maths prize



The new prize will honour Niels Abel.

**Oslo** Norwegian mathematicians have announced a new reward to honour successful researchers. One hundred years after the idea was first mooted, the country has established a lucrative maths prize that it hopes will rival the Nobel Foundation's. Norwegian Prime Minister Jens

Stoltenberg announced his government's decision to create a fund of Nkr200 million (US\$22 million) to support the annual prize last month. The first prize of Nkr5 million will be awarded in 2003. The Fields Medal, currently the most prestigious maths prize, is only awarded every four years and comes with substantially less money.

A Norwegian maths prize was first proposed in 1902 by King Oscar II of Sweden and Norway. The award will be known as the Abel Prize, after mathematician Niels Henrik Abel — 2002 is the bicentenary of his birth.

## Political action committee seeks cash from scientists

**Washington** US scientists and institutions are likely to be tapped for political contributions after the setting up of the first political action committee (PAC) devoted to science.

SciPac, as the new organization is called, will politely ask supporters of science to contribute \$250–1,000 to help Republican candidates for Congress. PACs have emerged in recent years as way for special-interest groups to feed money — and influence — into American politics.

Vernon Ehlers (Republican, Michigan), a former nuclear physicist, established the committee late last month, and will select its beneficiaries. "It'll give more visibility to science," he says. "And it will make more of the fact that Republicans have been very friendly to science: most scientists harbour the false impression that the Democrats are more friendly." Democrats decried Ehlers' move: one called it "an outrageous effort to use science for party-political purposes".

## Web gateway offers biological integration

**Paris** E-BioSci, a non-profit website intended to integrate disparate literature with gene, protein and other biological

databases, was officially launched last week.

The database was initiated by the European Molecular Biology Organization in response to US plans to create PubMed Central, a free website for life-science papers (see *Nature* **401**, 413; 1999). It received 2.4 million euros (US\$2.16 million) in funding from the European Commission earlier this year.

E-BioSci is requesting publishers to submit the full text, and not just abstracts, of the articles they hold, to allow cross-searches of the websites of participating publishers (see *Nature* **413**, 1–3; 1999). Nature Publishing Group (NPG) has already agreed to submit journal content to E-BioSci. "We look forward to working with E-BioSci to develop search and other functionality across the full text of papers published in NPG journals and those of other collaborating publishers," says Annette Thomas, the group's managing director.

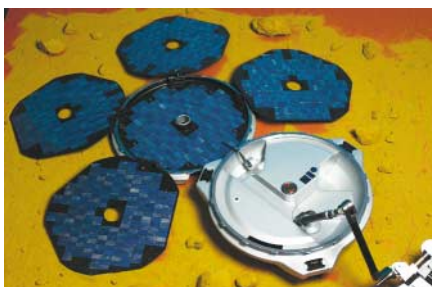
♦ <http://www.e-biosci.org>

## National projects to benefit from space agency ruling

**Stockholm** The European Space Agency (ESA) is to formalize its recent practice of partially funding national space-science projects.

ESA's constitution has always allowed it to support science in national agencies, but it





**Top dog:** British-built Beagle 2 was one of the first national projects to receive ESA funding.

only started to do so in the past couple of years. In that time, it has supported the Mars lander Beagle 2, which had been developed by British scientists for flight on ESA's 2003 Mars Express mission, and two French-led missions, the planet-finding Corot mission and the fundamental-physics mission Microscope.

A meeting of ESA's Science Programme Committee in Stockholm last week agreed to set up a task force to establish ground rules for selecting national missions that are worthy of support. "We will switch from a first-come-first-served approach to one where rules of engagement will be clear and projects can compete with each other," says Giovanni Bignami, vice-chairman of the committee.

## Millions for mouse mutant database

**Washington** The mouse genome database at Jackson Laboratory in Bar Harbor, Maine, has been given a boost of \$35 million over five years by the National Institutes of Health.

The laboratory is the main source of mutant mouse strains worldwide. Its database includes detailed genomic data on the strains, as well as information on those whose particular characteristics could make them useful in studying different diseases. The grant will be used to help the lab to develop the database, add additional data and improve its graphical interface.

The grant comes at a time when two other initiatives, the European Mouse Mutant Archive and the European Molecular Biology Laboratory, both of which are in Monterotondo near Rome, are also starting to promote the use of the mouse as a disease model.

## Cancer centre prepares to tackle conflicts of interest

**San Diego** Stringent rules that would block researchers from participating in clinical trials in which they have an economic

interest are being proposed at the Fred Hutchinson Cancer Research Center in Seattle, Washington.

The new proposals were generated by an external panel that was established earlier this year by the centre's trustees, after reports of problems with clinical trials in which participating doctors and the centre itself had possible conflicts of interest. The centre's board will now consider adopting the panel's recommendations, which include ensuring the disclosure of any institutional conflicts to patients, increasing numbers of staff for monitoring clinical trials, and better reporting of unexpected incidents.

If the rules were adopted, the centre would have some of the toughest conflict-of-interest standards in the United States. Many US facilities allow researchers to have a financial interest in drugs or devices under study, although some institutions increase monitoring in such cases.

♦ <http://fhcrc.org>

**Erratum** The photograph that accompanied last week's News story "Liver tumours temper hopes for gene-therapy technique" (*Nature* **413**, 9; 2001) incorrectly showed adenovirus particles rather than adeno-associated viruses. The technique in question uses the latter for gene delivery.



## Peers under pressure

As journal editors and scientists meet this week to discuss peer review, Rex Dalton considers what happens when competitive pressures disrupt the process, and examines measures designed to keep the system straight.

In February 1987, just months after the discovery of high-temperature superconductivity in materials known as cuprates, labs worldwide were locked in heated competition to identify new members of the family. That month, in two papers submitted to *Physical Review Letters*, researchers led by Paul Chu of the University of Houston and Maw-Kuen Wu of the University of Alabama in Birmingham described a new superconducting material as “ytterbium barium copper oxide”. But with the manuscripts on the verge of publication, the authors changed the name of the material: the element “ytterbium” became “yttrium”.

Many materials scientists suspect the error was deliberate, designed to protect the discovery during the peer-review process. Chu originally blamed a typographical error, and today refuses to comment. But it is clear

that he and Wu had reason to worry. “There is no doubt the manuscript was leaked during peer review,” says Paul Grant, a superconductivity researcher at the Electric Power Research Institute in Palo Alto, California, who has followed the story closely. While the paper was under review, says Grant, researchers from at

least two other groups contacted Chu to tell him that the material with the erroneous element did not work as a superconductor.

Speak to researchers in certain fields, and tales about abuses of peer review start creeping from the woodwork. Some scientists complain of manuscripts being stalled in review until similar findings emerge in another journal. Was the reviewer an author of the second paper who failed to declare this conflict? Did he or she deliberately delay the manuscript so as to publish first? Had the competing paper already been submitted to another journal, or could the work have been conducted, submitted and published after reading the manuscript? Other researchers worry about reviewers who are consultants to companies and who could pass on information that should remain confidential. Might such instances help companies win the race to market new discoveries?

### Articles of faith

These issues will be debated this week, as journal editors, scientists and other interested parties meet in Barcelona for the Fourth International Congress on Peer Review in Biomedical Publication. The conference — held roughly every four years — is the brainchild of Drummond Rennie, a professor of medicine at the University of California, San Francisco, and a deputy editor of the *Journal of the American Medical Association*.

Rennie is a strong supporter of the value

of peer review, and *Nature's* interviews with other journal editors and researchers reveal that most agree with him. By and large, they argue, peer review is a mutually beneficial system that provides effective quality control in publication and grant awards. But although instances of abuse may be rare, the process is not problem-free.

In some areas, such as geology — where important findings often depend on years of meticulous fieldwork — complaints are few and far between. But the potential for abuse is greater in fast-moving fields such as molecular biology, in which it is sometimes possible to read a manuscript and replicate the work in a matter of days. Where problems arise, the intense drive of individual scientists to further their standing is most often to blame. But commercial and, in rare cases, political pressures (see ‘Igniting an unholy row’, page 104) can also come into play.

Reviewers who directly misuse the information contained in manuscripts or grant applications are only part of the problem. Many reviewers admit to discussing papers or grant applications that they are sent for review with their colleagues, which increases the chance that someone might abuse that privileged information.

In late 1996, for instance, molecular biologist Carolyn Price, then at the University of Nebraska at Lincoln, was reading through grant applications in molecular cytology for the National Institutes of Health (NIH) when one application looked suspiciously familiar. She pulled out one of her



**Drummond Rennie:** striving to improve a valued system.

own grant applications, which she had submitted earlier that year to a state funding agency, and found that sections had been plagiarized.

"It was bizarre — a twilight-zone experience," says Price, now at the University of Cincinnati in Ohio. A subsequent probe revealed that a Los Angeles-based reviewer examining Price's application had shared the document in confidence with biochemist Ashraf Imam at the University of Southern California (USC). The Office of Research Integrity, the watchdog body that investigates allegations of misconduct surrounding NIH grants, concluded that Imam had plagiarized material from Price's application. In 1998, he agreed to be debarred from receiving NIH funds for three years. Imam now works on gene therapy at the Huntington Medical Research Institutes in Pasadena, California. Clive Taylor, professor of pathology at USC and Imam's boss at the time, attributes the event to "the enormous pressures on non-tenured faculty".

John Hardy, a neuroscientist who is this month moving from the Mayo Clinic in Jacksonville, Florida, to the National Institute on Aging in Bethesda, Maryland, says that he sometimes discusses manuscripts sent to him for review with lab colleagues — with the proviso that they cannot reveal the material to anyone else or use it for a publication or grant application. Hardy argues that this is common among researchers in his field, and believes the issue needs to be recognized and debated. "I don't believe people when they tell me they don't talk to people about manuscripts they receive for review," he says.

But some researchers would like to see the

## It is often difficult to distinguish between genuine complaints and paranoia.

practice stamped out. "It is terrible," says Randy Schekman, a molecular biologist at the University of California, Berkeley, and an editor of the *Journal of Cell Biology*. "I know it occurs, but I hope not widely. I think it is a breach of confidentiality that really compromises the process," he says.

Despite Schekman's reservations, most scientists questioned by *Nature* argue that limited discussion of papers is acceptable and should improve the quality of review. Some journals do issue clear guidelines, however. *Nature*, for example, asks its reviewers not to discuss confidential manuscripts with colleagues unless this is necessary to assess its merits — and where this is the case to provide the names of those involved.

### Commercial breaks

When a paper contains commercially sensitive information, breaches of confidentiality are a particular concern. Many researchers act as consultants to companies, and once papers under review start to be discussed, there is always the chance that sensitive information will reach the ears of someone with a commercial interest in it. Many researchers are certain that this happens, but such cases are notoriously hard to prove.

In 1991, a genetics paper submitted to *Nature* was at the heart of one such murky incident. Written by a group at a British university, the paper contained details of a discovery likely to be of interest to many biotechnology companies. One author of the paper is convinced that a company knew of the discovery before the paper was published.

His suspicions were first aroused about two weeks before publication, when he learned that a *Newsweek* reporter had been asking other researchers in the field about the paper. Then, just two days after publication, the group received a visit from a senior representative of a US-based biotechnology firm who wanted to license the rights to develop new therapies based on the discovery. But despite the author's suspicions of a breach of confidentiality, the facts of the case provide no 'smoking gun'.

*Nature's* interviews also reveal that it is often difficult to distinguish between genuine complaints and paranoia. Many editors can relate tales of anguished authors convinced that a reviewer had abused the process. Such arguments often surround negative reports perceived by the author to



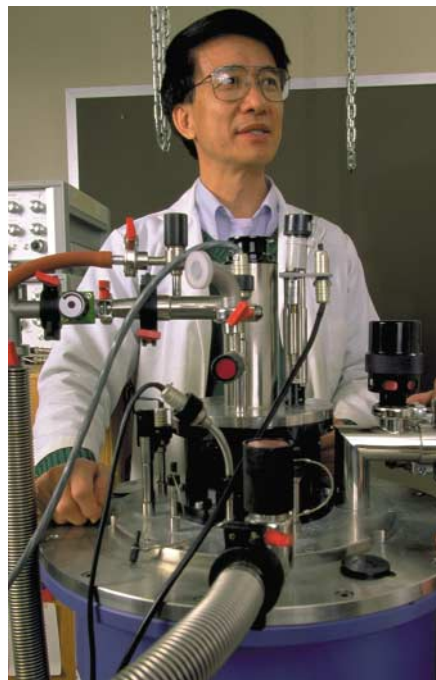
No doubt: Paul Grant (right) is certain that leaks occur, particularly in highly competitive fields.

have come from a particular rival intent on blocking the paper's publication — but who had not, in fact, been sent the manuscript. "An author will suggest who to send the manuscript to," says Schekman. The journal may do so, but if that reviewer delivers a negative report, the author may assume wrongly that it came from his or her rival. "Then they howl, because they don't like the response," Schekman says.

But despite the problems thrown up by peer review, no serious alternative has yet been proposed. "It is easy to say the system is flawed; it is harder to say how to improve it," says Ronald McKay, a stem-cell researcher at the National Institute of Neurological Disorders and Stroke in Bethesda, Maryland.

One tweak to the process — asking reviewers to sign their reviews — has been experimented with. The idea is that, if reviewers are obliged to identify themselves, it will improve transparency and discourage anyone who might be tempted to abuse the process under the cloak of anonymity. Rennie is a particular enthusiast for this approach. "This is the only credible, worthwhile, transparent and honest system," he says. "I've made that passionate plea, but the majority hasn't gone along with it."

Indeed, when the idea was first mooted many years ago, some editors were openly hostile. In the 1960s, when David Sackett, now professor emeritus at McMaster University in Hamilton, Ontario, started signing his reviews, he was removed from the review boards of journals such as *The*



Paul Chu was contacted about his paper on superconductors while it was still under review.

UNIV. HOUSTON

DETROIT EDISON



## Igniting an unholy row



**A manuscript critical of the National Ignition Facility was widely circulated during review.**

researcher at the Natural Resources Defense Council, a Washington-based environmental group.

*Science* discounted Moses's review, but still rejected the manuscript. The rejection was based on other reviews and "the complexities of the issue", noted editor-in-chief Donald Kennedy in a letter to Paine and Bodner, adding that "all of us are appalled at the behaviour of the referee for violating the confidentiality instructions". Kennedy says that Moses and other senior figures at Lawrence Livermore received a "tough" letter from *Science*.

Moses declined to be interviewed, but Kennedy says that Moses apologized, arguing that he had not realized that the manuscript was meant to be confidential. "I received a promise it wouldn't happen again," says Kennedy. Ultimately, an edited version of the manuscript appeared in *Nature* (S. Bodner & C. Paine *Nature* 407, 129–130; 2000).

Usually, abuses of peer review appear to be driven by academic or commercial rivalries. But political pressures can also play a role, as physicists Stephen Bodner and Christopher Paine can attest.

Last year, Bodner and Paine submitted an opinion piece to *Science* criticizing the lack of independent peer reviews of proposals for the National Ignition Facility (NIF), a giant laser project at the Lawrence Livermore National Laboratory in California. At the time, the project's cost overruns were attracting the attention of the US Congress. And in this politically charged atmosphere, the piece was sent to reviewers including Edward Moses, NIF's programme manager.

Moses penned a scathing review recommending rejection, which his office then sent out to dozens of other physicists. The authors discovered the leak when Bodner, who retired in 1999 from his post as head of the laser-fusion programme at the Naval Research Laboratory in Washington, received a copy of the review as it bounced through physics list-servers on the Internet. "We were outraged," says Paine, a senior

Where it seems that a reviewer has abused the process, the most common course of action is to challenge the individual concerned, and if he or she cannot convince the editor of their innocence, for the journal to cease using that person as a reviewer. *Science's* editor-in-chief, Donald Kennedy, says that, in particularly egregious cases, a journal might refuse to consider an offender's own manuscripts for publication.

### Penalty clause

Other editors have, on occasion, referred accusations of misconduct during peer review to the accused individual's host institution. Nicholas Cozzarelli, professor of biochemistry and molecular biology at the University of California, Berkeley, and editor-in-chief of the *Proceedings of the National Academy of Sciences*, says he has found this to be the most effective policy. But he says that taking such action requires an accuser to make a formal complaint — which some aggrieved manuscript authors are loath to do for fear of attracting a lawsuit from the accused.

Philip Campbell, editor of *Nature*, says that allegations of abuse during peer review are dealt with on a case-by-case basis. In addition to the sanctions listed above, *Nature* may also write an article about clear instances of misconduct. He acknowledges that it is rare that one can be sure enough to take the stronger sanctions. "But some employers, in becoming tougher against misconduct, are encouraging us to report even suspicion of bad behaviour in peer review. We'd be very cautious about that."

Many editors also argue that procedures to minimize and identify abuses are more important than sanctions against those who offend. Journals usually demand that reviewers declare any conflicts of interest that might influence their ability to provide an unbiased assessment of a paper. Simple procedures such as acceding to authors' requests not to send manuscripts to particular competitors also help to avoid problems. And if manuscripts are routinely sent to two or three different reviewers, suspicions should soon arise if one reviewer raises spurious objections to a manuscript in an attempt to block a rival's publication.

The Barcelona meeting may well throw up other suggestions for improving the system. But despite its flaws, peer review will continue to form the bedrock of the scientific enterprise. For every allegation of abuse, there are countless more papers that are improved by helpful suggestions made by reviewers. "I've seen the amazingly wonderful things peer review does for authors," says Rennie. "I would hate to give it up."

**Rex Dalton is *Nature's* West Coast US correspondent.**

Fourth International Congress on Peer Review in Biomedical Publication

♦ <http://www.ama-assn.org/public/peer/peerhome.htm>

LLINE

### ► *New England Journal of Medicine*.

Attitudes have softened since then, but some editors remain convinced that the practice has its problems. "I personally discourage it," says Stephen Lisberger, a neurophysiologist at the University of California, San Francisco, and an editor for the *Journal of Neuroscience*. "It encourages an author to bypass the editorial process and correspond with the reviewer. Then you get negotiation and compromise between the author and reviewer." To minimize such problems, some journals — including *Nature* — only reveal the names of reviewers who want to be identified in the final round of review.

### Identity crisis

Lisberger also recalls an experiment in the 1980s at the *Journal of Neurophysiology* in which reviewers were encouraged to sign their reviews. "We found reviewers signed their positive reviews, but not their negative ones," he says. "It didn't change the back-biting, or complaints that so-and-so was blocking publication." The journal eventually dropped the policy.

Other scientists and editors argue that the Internet has helped to reduce abuses. One common fear is that reviewers may be deliberately slow in responding if they wish to delay a rival's paper. But instantaneous com-

munication by e-mail has greatly reduced the scope for reviewers to blame non-delivery of manuscripts and reports for such tardiness. Also, the availability of literature online means that researchers who plagiarize information gained during peer review are more likely to be found out. "You can scan journals so much quicker now," says Ralph Yount, a biochemist at Washington State University in Pullman and former president of the Federation of American Societies for Experimental Biology. "People can't steal things and publish in an obscure journal somewhere."

Such abuses may be relatively rare, but they are a serious business for the researchers on the receiving end. So what happens in such instances? As the Imam case shows, blatant abuses such as plagiarism of grant applications can lead to tough sanctions. But journal editors say they are rarely in a position to act as judge and jury when suspicions arise. Often, the facts are far from clear. For example, a reviewer might agree to review a paper on being sent its abstract, and then sit on the full manuscript for several weeks before returning it, declining to act as a reviewer because he or she has a competing manuscript under review with another publication. Does that represent a cynical attempt to delay a rival's publication? Or is it an innocent slip-up by a busy scientist?



# Keeping up with the Joneses

Massive financial endowments allow the top US universities to offer the best salaries and conditions in the academic world. David Adam asks how their British counterparts can close the gap.



**Rich heritage: but Cambridge (above) manages on a fraction of Harvard's (right) budget.**

It was the sort of comment that any British university researcher could have made over a cup of coffee. But when Neil Rudenstine, the outgoing president of Harvard University, told a seminar at his university earlier this year that Britain's higher education system was going from "a disaster to a nightmare", his remarks did not go unnoticed. Unfortunately for Rudenstine, they found their way into the British press and provoked outraged complaints.

Rudenstine may not be thanked for pointing it out, but many within Britain's universities agree with his sentiments. Britain's top institutions produce high-quality research — the government's Office of Science and Technology estimates that British researchers publish 8% of the world's science papers with only 4.7% of the world's research investment —

but they lag behind their wealthy US rivals when it comes to salaries and research facilities. Universities in most other developed countries suffer from a similar wealth gap. But in Britain, used to trading on a tradition of excellence founded by Oxford and Cambridge, concerns about universities' ability to compete are particularly acute.

"We punch above our weight, but we hang on by a thread," says Richard Sykes, non-

executive chairman of the drugs company GlaxoSmithKline, who took over as rector of London's Imperial College earlier this year. Unless the gap is closed, many fear that British universities will lose out in the competition for the best students and researchers. And with the government unlikely to provide enough extra investment, a variety of different funding models are being mooted.

The difference between funding in Britain and the United States is significant. Private US universities such as Harvard get the bulk of their income from tuition fees and the return on their endowment funds. Harvard's \$19-billion endowment, the largest in the United States, yields \$500 million each year. Yale, Princeton and Stanford each have endowments of over \$6 billion. Top public institutions, such as the University of California, Berkeley, benefit from generous state and federal funding, which allows them to compete with their private rivals.

## Poor relations

For the private US universities, much of this wealth comes from donations from former students, a trend almost unknown in Britain. Oxford and Cambridge — the wealthiest British universities by a considerable margin — have respective endowments of around £500 million (US\$723 million) and £650 million. Imperial College and University College London — two of the other universities that are most able to com-

pete on a world stage — both have endowments of less than £100 million.

Top British universities can still fight their corner in generating income from government, charities and industry to fund research projects. Oxford's external research income was about £130 million during the past year, which compares favourably with that of US giants such as Harvard, says David Holmes, the university's registrar.

But researchers at US institutions earn considerably more than their British counterparts, and have much better provision for research equipment. Differences in grading structures and funding arrangements make direct comparisons difficult, but this year, the top of the university-lecturer pay scale in Britain was about £40,000 — towards the low end of what someone in an equivalent position in the United States could expect.

British universities have also been stretched to breaking point by a rapid, but underfunded, expansion in student numbers over the past 15 years. And bringing the infrastructure back up to scratch, as well as making salaries internationally competitive, will not be cheap.



**Neil Rudenstine had a nightmare vision for Britain.**

AP

PA/HARVARD



A recent study by Universities UK, the London-based representative body for British universities, estimated that, in terms of teaching requirements, the higher education system is underfunded by around £900 million per year, and that a similar sum is needed on the research side. "We have managed to stay competitive so far, but the present situation is unsustainable and needs urgent review," says Diana Warwick, the organization's chief executive.

Over the next two years, the government is providing £1 billion to improve research infrastructure and around £280 million to improve pay and retain key researchers. But with even the most optimistic observers accepting that the government will not commit to a long-term increase in university funding, the search is on for alternatives.

## Pay as you learn

Allowing market forces to dictate the price of places on different courses at different institutions is one suggestion. At present, the government forbids any university to charge more than £1,075 per year in tuition fees for full-time undergraduates. Deregulating the system could benefit the top universities, who could use the high demand for places to push up tuition fees.

But such a plan would be controversial. The introduction of tuition fees in 1998 increased concerns that the cost of higher education is deterring applicants from low-income families — the long-standing system of maintenance grants for students was already being phased out in favour of loans. The government currently helps poorer students with tuition fees, but it may be unwilling to meet the full cost of deregulated fees.

"There is evidence from the United States

that the high cost of university education deters applicants from low-income and socially excluded groups," says Claire Callender, a social-policy researcher at South Bank University in London. Worries over this issue have already led the Scottish parliament to scrap tuition fees for Scottish students attending the country's universities.

One alternative is to make students pay for their courses once they graduate and start earning a salary above a certain level. By setting the contributions high enough, university income could be raised without increasing direct investment from the government. But payments could only be secured from new graduates. In the meantime, extra government money would be needed.

A third option — and the one favoured by

the opposition Conservative party — is to free universities from some state control by endowing them with the proceeds from future state-privatization programmes. Funding for research would still come from the government, but teaching grants, which currently stand at several tens of millions of pounds per year, would be replaced by the interest on a university's endowment.

An independent source of income would appeal to many universities, but the funds required are huge. Universities UK estimates that £24 billion would be needed to create endowments to cover the current teaching income of the 12 universities with the largest government-funded research programmes.

Of these options, a mixture of deregulation and endowment schemes is likely to appeal to the top British institutions. And regardless of whatever changes the government may choose to make, reports suggest that some universities — including Oxford and Cambridge — are already making plans to abandon state funding and move towards a US model based on building up an endowment and setting their own tuition fees.

The subject is "on the agenda nationally", Holmes admits, but the situation is a complicated one. "We would welcome as much financial independence as we can achieve without a doubt," he says. "But we're highly constrained by government policy at present and it's a highly political question."

## Two-tier system

The threat of deregulation poses a problem for the British government, as any break-away group charging higher tuition fees would effectively create a two-tier system: a major headache for a government claiming to be committed to widening access to higher education for people from less privileged backgrounds. So far, the government looks unlikely to raise the tuition fees it imposed, publicly stating that they will remain capped for the duration of this parliament.

But a restructuring may be inevitable. There are 111 publicly funded universities in Britain, each competing for funds and students. According to Sykes, only a few should be expected to compete on an international stage. "There is no question you are going to see a dichotomy," he says.

Such comments might be expected from the head of an elite institution, but some others outside the system agree that more diversification is needed. "Someone who is a world-class teacher of the most able students



Top UK universities such as Oxford (left) and Imperial need to find more money.



and a researcher working in a group with a global reputation can reasonably expect greater remuneration," says Peter Cotgreave, director of pressure group Save British Science.

But not everyone thinks further diversification within the sector is a

good idea. According to Callender, deregulation of tuition fees could reinforce the social divisions within universities. "Students from low-income families are already less likely to apply to top British institutions," she says. "This could make the system even more elitist." Callender suggests that a better funded and more extensive version of the current bursary scheme, perhaps paid for by employers and industry as well as the government, could go some way to addressing the problem. The United States already has a well established bursary system, which allows many students from low-income families to attend private institutions such as Harvard.

Rob Copeland, education policy research officer for the Association of University Teachers, the trade union for British university lecturers, points out that there is already some flexibility in both funding and salaries. Links with business are rewarded financially and top universities can pay some staff higher salaries through localized pay scales.

Avoiding charges of elitism while allowing Britain's universities to compete with their US rivals looks likely to prove a difficult juggling act for the British government. Unless other alternatives are pursued, the top institutions may take the matter into their own hands. The British system may not yet have turned from a disaster to a nightmare, but it does seem stuck between a rock and a hard place.

David Adam is a News and Features writer for Nature.



Diana Warwick believes an urgent review is needed.



# Investment is the best cure for inbreeding

Lack of funding and difficulty in setting up new labs encourage researchers to stay put.

Sir — M. Soler in Correspondence (“How inbreeding affects productivity in Europe”, *Nature* **411**, 132; 2001) quantifies endogamy, or “inbreeding”, in 14 European countries as the percentage of staff trained at the same university. His results are of great value as a basis for political decisions on university reform. He discovered a huge variation in endogamy, from 1% in Germany to 88% in Spain and 91% in Portugal. To uncover the causes, I have analysed some economic and demographic data published in the *Eurostat Yearbook 2001* (Eurostat, Luxembourg, 2001; <http://europa.eu.int/comm/eurostat/Public>).

The percentage of GDP invested in research and development is significantly negatively correlated with endogamy ( $r_s = -0.75$ ,  $P = 0.0018$ ). In terms of annual *per capita* investment in R&D, Portugal (67 euros; US\$60.6) and Spain (116 euros) invest the least, and Sweden (936 euros) and Switzerland (935 euros) the most, with the remaining countries showing intermediate values. This indicates that endogamy is a consequence of poor investment policies.

Because there has been much discussion about Spain, I will refer to it here as an example. Spain is the second-highest country in Soler's endogamy index, but has the lowest cost per paper appearing in the *Science Citation Index (SCI)*. This leads one to ask whether endogamy is indeed negative for the R&D system?

To answer this question, I calculated three indexes of productivity. One is simply the number of *SCI* papers *per capita*, which is negatively correlated with endogamy ( $r_s = -0.73$ ,  $P = 0.003$ ). The second is the number of *SCI* papers per researcher. I obtained the number of researchers (including technicians) per country (except the United Kingdom and Germany) from the World Bank web page (<http://wbln0018.worldbank.org/psd/complete.nsf/f14ea5988b0e0ec7f852564900068cbfd?OpenView>). This second index fails to show significant correlation with endogamy ( $r_s = -0.17$ ,  $P = 0.59$ ). My third index is the cost per paper appearing in the *SCI*, calculated as the quotient between the absolute amount of money invested in R&D and the number of papers appearing in the *SCI*. This index also failed to show significant correlation with endogamy ( $r_s = -0.37$ ,  $P = 0.19$ ).

Taken as a whole, these results could reflect the unequal numbers of researchers per 10,000 inhabitants, varying from 1 in Spain and Portugal to 6 in Sweden, and

perhaps the different effort demanded of researchers in different countries.

In my opinion, Spanish endogamy is caused by low investment in R&D (0.9% of GDP in 1999). Spanish universities lack research technicians, so researchers have to do everything: collecting materials in the field, preparing samples for analysis, performing all laboratory techniques, and so on, as well as teaching. Spanish *SCI* papers are thus low-cost, but single researchers have very difficult lives, impelling us to integrate into groups for support — a process which is also encouraged by the policies of regional governments. On average, we need 10 or more years to create a research group with acceptable productivity.

Moving is very difficult for many

reasons, some discussed by previous correspondents. Those who do move have to start a new lab from scratch with no equipment and without colleagues from the previous lab. The logical consequence of these factors is reluctance to move, and hence endogamy.

I agree with Soler and other correspondents that this situation needs to be corrected. My analysis shows that the best and simplest solution for Spain is significantly increased investment in R&D, which would avoid the causes of endogamy. But I am afraid that the government's proposed reform is opting for the cheapest solution, not the best one.

**Juan Pedro M. Camacho**

Department of Genetics, Faculty of Science, University of Granada, E-18071 Granada, Spain

## Taxonomy is small, but it has its citation classics

Sir — F. T. Krell in Correspondence (“Impact factors aren't relevant to taxonomy”, *Nature* **405**, 507–508, 2000) suggests that taxonomy is subject to different “regularities” from other fields. He uses the allegedly small number of entomology, biodiversity and taxonomy journals covered in the *Science Citation Index (SCI)* as the reason for low impact factors. But he provides only anecdotal data on the size of this literature. How many articles are published by the low-impact journals at the Natural History Museum?

Krell refers to Bradford's law of scattering, but he needs instead to provide data to show that entomology–taxonomy diversity is somehow different from other fields. Bradford's law simply suggests that if there are, say, 1,000 journals in a field, then one-third of the papers are to be found in each of three zones containing about 10, 100 and 1,000 journals, respectively.

A critical element in determining the impact factor of a field is not the number of papers it publishes, but the citation density of the average paper and the half-life of the references cited. Adding more journals to the *SCI* database would not increase the impact because the increased number of cited references would have to be shared by more published papers.

The data on the 65 entomology journals covered in the *SCI* indicate that their impact factors are not significantly

different from other fields with long half-lives (see the Institute for Scientific Information's *Journal Citation Reports*, <http://www.isinet.com/isi/products/citation/jcr/?version=1>). However, the size of a field does affect the number of super-cited papers that will be published.

Krell states: “Qualified referees must evaluate the scientific work itself.” But this is equally true of any field. Why shouldn't such evaluation be supplemented by citation analysis? The question is whether the referee has any basis for comparing articles, authors or journals across a wider horizon. Taxonomy, small as it may be, is not without its ‘citation classics’, as the work of R. Sokal, E. O. Wilson and others demonstrates. Their work is cited by thousands of papers covered in the *SCI*, by taxonomists and by other scientists.

**Eugene Garfield**

Institute for Scientific Information, 3501 Market Street, Philadelphia, Pennsylvania 19104, USA

## Ethical link between IVF and stem-cell research

Sir — I fully support your editorial “The meaning of life” (*Nature* **412**, 255; 2001), in which you argue that individual life does not begin with fertilization of the egg and that it thus should not be used as an argument against creating embryonic stem cells, which requires the destruction of fertilized embryos. Hubert Markl's Commentary “Research doesn't denigrate humanity” (*Nature* **412**, 479–480; 2001) makes a similar point.

However, I regret that you missed the opportunity to point out that human reproduction by *in vitro* fertilization (IVF) also involves the fertilization of the egg and the early development of the embryo, and that large numbers of such embryos are destroyed.

There is thus no ethical difference between IVF and creating embryonic stem cells, as both require the creation and destruction of embryos. One can be, for religious reasons, against both, but not rationally against one and not the other. IVF has been of enormous value and so too will stem cells.

**Lewis Wolpert**

Department of Anatomy and Developmental Biology, University College London, Gower Street, London WC1E 6BT, UK

## Science archives should remain in public hands

Sir — We would like to correct any impression of neglect of Britain's rich scientific archive heritage that might have been given by the News feature "The History Man", about the US private collector Jeremy Norman (*Nature* **411**, 732; 2001). On the contrary, the United Kingdom is fortunate in having universities and national institutions that show an active interest in collecting science archives.

The Royal Society and the Wellcome Trust, for example, have supported the preservation and cataloguing of such materials over many years. The Wellcome Trust has recently established a Research Resources in Medical History scheme, with £1 million (\$US1.4 million) for the year 2001–2002 to support important documentary collections.

Archives in universities, including contemporary science archives, have benefited from major funding programmes run by the Higher Education Funding Councils. A recent award from the UK Heritage Lottery Fund will allow a group of science institutions to mount a large number of catalogues of scientists' archives on the web as part of Access to Archives, a vast online catalogue at <http://www.a2a.pro.gov.uk>.

Many important personal scientific papers are held in libraries and repositories in the United Kingdom, some having been catalogued by the National Cataloguing Unit for the Archives of Contemporary Scientists (NCUACS). This work is supported by several scientific societies, trusts and foundations, preserving a significant part of contemporary British science and biomedicine in a major collaborative effort.

For the long-term benefit, such papers are best housed in properly resourced public repositories in their country of origin, rather than in private hands. Archivists, who will always have to struggle to maintain their budgets in a competitive world, will be greatly helped by widespread recognition of that basic principle.

**Peter Harper\*, Julia Sheppard†**

\*NCUACS, University Library, University of Bath, Claverton Down, Bath BA2 7AY, UK

†Wellcome Library, 183 Euston Road, London NW1 2BE, UK

## Nature's laws revealed in rhyming couplets

Sir — I would like to add to Fleming Carswell's interesting Correspondence (*Nature* **411**, 885; 2001) about the relevance of poetry to scientists today by mentioning a few of the scientists who in the past published their work as poetry.

Charles Darwin's grandfather Erasmus Darwin (1731–1802), for example, wrote up some of his own evolutionary and other theories in Popean couplets, perhaps best known in *The Temple of Nature*:

*Organic life beneath the shoreless waves  
Was born and nurs'd in Ocean's pearly caves;  
First forms minute, unseen by spheric glass,  
Move on the mud, or pierce the watery mass.*

In the same poem he describes 'the Maiden Truffle' as an example of reproduction without a sexual partner:

*So the lone Truffle, lodged beneath the earth,  
Shoots from paternal stems the tuberous birth.  
No stamen-males ascend, and breathe above,  
No seed-born offspring lives by female love.*

These influenced not only his grandson but also the English romantic poets, to such an extent that Samuel Coleridge used the term "Darwinizing" to describe such poetic theorizing.

Alexander Pope (1688–1744) himself asked a cosmological question in his *Essay on Man*, Epistle 1, which is only now being answered:

*Observe how System into System runs.  
What other Planets circle other Suns?*

Later in the same epistle, he seemed to lay the foundation of statistics:

*All Nature is but Art unknown to thee;  
All Chance, Direction which thou canst not see.*

Pope also wrote in his *Epitaph for Sir Isaac Newton* (perhaps anticipating the need for this journal?):

*Nature, and Nature's laws lay hid in night  
God said, Let Newton be! and all was light.*

to which J. C. Squire (1884–1958) in his *Epigrams* replied:

*It did not last: the Devil howling "Ho!  
Let Einstein be!", restored the status quo.*

The advice Pope gave in his *Essay on Criticism* could apply to *Nature's* referees:  
*Let such teach others who themselves excel  
And censure freely who have written well.*

Modern poets offer yet more of the stimulating and enjoyable reasons why, as Carswell states, scientists should "bother about poetry". Not least among these are the poetry of Edwin Morgan (to be found, for example, in *Stargate*, Third Eye Centre, Glasgow, 1979); Miroslav Holub (for example, *Vanishing Lung Syndrome*, Faber and Faber, London, 1990); and Ronald Duncan (for example, *Man: The Complete Cantos*, Rebel Press, London, 1970).

**N. C. Craig Sharp**

Department of Sport Sciences, Brunel University, Borough Road, Isleworth, Middlesex TW7 5DU, UK

## Enthusiasm ran ahead of discoveries still to come

Sir — In your excellent News story "Epidemiology gains an ally in bioweapons surveillance project" (*Nature* **411**, 228; 2001), summarizing the rapid syndrome validation project (RSVP) for early 'syndrome-based' epidemiological reporting, you faithfully reported my statement that "This rapid reporting already appears to have this year averted two outbreaks of hepatitis A" and that antiviral-drug prescription frequency may have been altered.

This information was incorrect. I based these statements on early, anecdotal evidence that I cannot fully support. I should have checked my sources more carefully. I regret the error, and take full responsibility for it. All other aspects of the News story are accurate.

I look forward to reporting, in a future paper, the benefits and downsides of the novel systematic approach to disease reporting that we are taking at RSVP, once we have accumulated a sufficiently large database. We are fortunate to have funding from the US Department of Energy to expand the project to multiple reporting sites in various clinical settings, including non-academic medical clinics and public-health clinics on both sides of the Mexico–US border.

**Alan P. Zelicoff**

Sandia National Laboratories, PO Box 5800, Albuquerque, New Mexico 87185, USA



# The 'new synthesis' vindicated

Sociobiology is here to stay and the debate now needs to move on.

**The Triumph of Sociobiology**  
by John Alcock  
Oxford University Press: 2001. 257 pp.  
£16.95, \$27.50

Laurent Keller

Sociobiology has a peculiar history. Most people acknowledge that natural selection has been important in shaping the evolution of organisms and their genomes. Few would also deny that reproductive success, which determines which genes are transmitted from one generation to the next, is influenced by how organisms behave. One would think, therefore, that studies of social behaviour from an evolutionary perspective should be uncontroversial. So why is there such a fuss about sociobiology?

The controversy started soon after the publication of Edward O. Wilson's classic book *Sociobiology: The New Synthesis* (Harvard University Press, 1975). Most of the book was devoted to the evolutionary relationship between an animal's behaviour and its environment, in particular the social environment. This approach was far from new, beginning with the publication in 1859 of *On the Origin of Species* by Charles Darwin. Following the work of other evolutionary ethologists, Wilson tried to explain apparently peculiar behaviours, such as how insect colonies with sterile workers evolved or why some female spiders eat their mates

after mating. Sociobiology is based on the principle that, over evolutionary time, natural selection has favoured genes that increase the survival and reproduction of organisms. As a result, extant organisms should generally behave so as to maximize their reproductive success. So far so good. But Wilson went one step further in asserting, in the final chapter of his book, that the same approach can be used to study human behaviour. And this ignited controversy.

In his excellent book, John Alcock analyses the history of this controversy. The most vehement critics of sociobiology were some of Wilson's colleagues at Harvard. They were highly unreceptive to the notion that an evolved human nature exists, fearing that this might be interpreted to mean that human behaviour cannot be changed and that sociobiology could be used to justify unpleasant features of human behaviour. But, as clearly explained by Alcock, sociobiology by no means provides an ideological foundation for endorsing unwanted behaviour such as racism, fascism and sexism. Yet it is true that some of Wilson's statements, in particular those pertaining to the existence of a human nature, did not take into account history and how various definitions of man's 'natural state' had been used in the past to justify non-egalitarian political and social systems. Fortunately, Wilson and most other sociobiologists have now become more



**Inexplicable behaviour: a female mantis neatly decapitates her partner after mating.**

careful about the possible implications of their writing, as exemplified by Wilson's more recent book *Consilience: The Unity of Knowledge* (Random House, 1999).

Alcock addresses numerous misconceptions about sociobiology. Some were deliberately introduced by critics. For example, sociobiologists are frequently portrayed as believing in a 'biological determinism' or 'genetic determinism' of behaviour. No sensible sociobiologist believes such a thing. There is no doubt that human behaviour is the product of complex interactions between numerous genes and the social environment. The straw man of a simple genetic determinism was raised by critics such as Stephen Jay Gould because it is easy to chastise sociobiologists for getting their genetics wrong.

Behavioural differences between men and women provide a good example of the combined effect of genes and social environment. It is quite possible that men and women might have maximized their fitness by behaving differently in the past. As a result, genetic, hormonal and other physiological differences might have evolved, inducing, for example, a higher tendency to promiscuity or more attempts to attain a high social status in men than in women. But



**Rutting season: seeing off a younger rival.**

YANN ARTHUS-BERTRAND/CORBIS

PASCAL GOETHELUCK/ARDEA LONDON

even if men and women have inclinations to behave differently, there are no fixed and irreversible differences, and, as Alcock points out, sensible sociobiologists do not believe there are. All human behaviour is influenced by culture and can be modified by education.

It is not certain, however, that the best course for the well-being of men and women is to refute possible intrinsic differences and implement educational rules aimed at making men adopt a more 'woman-like' behaviour, or vice versa. Rather, it might be more helpful to acknowledge any differences and to formulate a policy that prevents social discrimination of one sex over the other.

Alcock also addresses the criticism that sociobiologists tell 'just-so stories'. He correctly points out that evolutionary theory makes predictions that can be tested. For that reason, he argues, sociobiology is just as rigorous as any other scientific field. Here I would add a note of caution. It is very difficult to test whether human behaviour is adaptive (or was in the past), not all studies are solid and there are publication biases.

Take the example of facial and body symmetry. There is evidence that greater asymmetry reflects higher stress during development as the result of a less favourable environment and/or 'bad' genes. Some sociobiologists thus predicted that females should be more attracted by more symmetrical males because they would potentially be better fathers. These predictions are supported by some studies, primarily in birds and humans, but Richard Palmer showed that this is a good example of selective reporting. Studies that found a negative association between male asymmetry and reproductive success were more likely to be published in scientific journals (not to mention the general media) than those that did not. Sociobiology, and evolutionary biology as a whole, would greatly benefit if selective reporting could be prevented. The risk of just-so stories would also be much lower if formal replicative studies were performed more often.

I was surprised by the book's title, as many readers probably will be. But it soon becomes clear that Alcock is not implying that sociobiologists are all correct and their critics all wrong. Rather, his point is that a large body of work now shows that animal behaviour, and to some extent human behaviour, has been shaped by natural selection. Thus, the debate should no longer focus on the merit of sociobiology *per se*, but should move on to more interesting issues such as the study of interactions between genes, social environment and culture. An even more challenging task, in my view, will be to acknowledge that we are not all identical, free of the influence of our genes, culture and education, while ensuring that this does

not lead to social discrimination between ethnic groups, genders and individuals. ■  
*Laurent Keller is at the Institut d'Ecologie, Bâtiment de Biologie, Université de Lausanne, CH-1015 Lausanne, Switzerland.*

## Shedding light on a golden age

**Nearest Star: The Surprising Science of Our Sun**

by Leon Golub & Jay M. Pasachoff  
*Harvard University Press: 2001. 267 pp.*  
\$29.95, £20.50

John H. Thomas

Solar physics is in something of a golden age. Recent observational results from highly successful space missions have significantly altered our understanding of the Sun's outer atmosphere, its magnetic-activity cycle and its influence on the Earth and the near-space

environment. The new technique of helioseismology — the probing of the solar interior using observations of oscillations at the solar surface — has given us a much more complete and accurate picture of the Sun's internal structure and dynamics. Measurements of the flux of neutrinos from the Sun are forcing changes in our understanding of fundamental particle physics.

*Nearest Star* beautifully presents these and other recent advances for the general reader, while also giving a good historical perspective on our study of the Sun. The authors are especially well qualified to write a popular book on this topic. Leon Golub is an astrophysicist at the Harvard-Smithsonian Center for Astrophysics. He has carried out observations of the Sun using rockets and satellites in space for more than 30 years, most recently as an investigator for NASA's Transition Region and Coronal Explorer (TRACE) spacecraft. Jay Pasachoff is professor of astronomy at Williams College in Massachusetts and an experienced writer of astronomy textbooks, who has observed 31 solar eclipses.



## Explosive secrets of the Sun

Images of coronal mass ejections such as this one, obtained using the Soft X-ray Telescope aboard the *Yohkoh* spacecraft, are helping to explain how such sudden explosions occur. These events, which come about when plasma and magnetic fields are transiently ejected from the Sun's corona — the outermost region of the solar atmosphere — produce intense shock waves, accelerating vast quantities of energetic

particles. When directed at the Earth, coronal mass ejections can cause strong geomagnetic storms, disrupting communications. Magnetic changes that occur before this energy release may act as a warning of imminent ejections. More on this and many other aspects of the Sun can be found in *The Cambridge Encyclopedia of the Sun* by Kenneth R. Lang (Cambridge University Press, £29.95, \$49.95).



The science of solar physics has two distinct aspects, one emphasizing the astronomical side of the subject (the 'solar–stellar connection') and the other the relationships between the Sun and the Earth. Both aspects are well covered in this book. On the astronomical side, the authors vividly describe the evolution of the Sun, drawing on current knowledge of star formation and evolution. The story runs from the Sun's birth in a collapsing interstellar cloud, through its youth and long middle age as a main-sequence star, its passage through the red-giant phase and the ejection of a planetary nebula, and finally to its slow death as a cooling white dwarf formed from the solar core.

The text provides remarkably clear explanations of some quite complicated things, such as the Sun's sharp edge, limb darkening, the formation of absorption and emission lines in the solar spectrum and the three-dimensional, dynamic nature of the solar chromosphere. There are also clear descriptions of the workings of several solar instruments, including the spectroheliograph and the birefringent filter. Interesting historical nuggets enliven the text throughout.

The authors are naturally at their best in discussing their own research specialties, solar eclipses (Pasachoff) and the solar corona (Golub). They give the reader a good feel for the careful planning, risks, tension, joys and frustrations of a solar-eclipse expedition or a space experiment. Several remarkable TRACE images of the solar corona, showing the beautiful, fine structure created by the Sun's magnetic field, enhance the volume, as do some important results from the Solar and Heliospheric Observatory (SOHO) and other recent space missions.

The book does have a few shortcomings. In their discussion of helioseismology, the authors fail to point out the basic fact that the waves causing the oscillations are sound waves. Subsurface magnetic fields are detectable not, as the authors state, because the magnetic field decreases the propagation speed by lowering the local temperature, but rather because it increases the propagation speed directly through magnetic restoring forces. There is little or no discussion of recent advances in high-resolution observations from ground-based telescopes, such as the use of image-restoration techniques or adaptive optics. But these are minor quibbles; overall, the coverage is broad and complete and the level of accuracy very high.

A timely feature of the book is its excellent chapter (entitled "Fire and Ice") on the Sun's influence on the Earth's climate. This chapter contains a thoughtful discussion of the climate system and global warming and shows why we need to have a better understanding of the effects of solar variability on our climate before we can sort out the man-made effects.

With the publication of this book,

Harvard University Press continues a tradition of excellent books on the Sun for general audiences. *Nearest Star* is an up-to-date, authoritative and entertaining introduction to the Sun for the general reader. It represents popular science writing at its best — expert authors writing in a clear and lively style, without oversimplification, engaging the reader's creative thinking and imagination. ■

John H. Thomas is in the Departments of Mechanical Engineering and of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA.

## Techy in fuzzy clothing

**Operators and Promoters: The Story of Molecular Biology and Its Creators**

by Harrison Echols, edited by Carol A. Gross

University of California Press: 2001. 466 pp. \$65, £45

Horace Freeland Judson

Textbooks in the sciences are typically ahistorical — hardly a word about how the science was created. Scientists and their textbook publishers are sceptical of a historical approach. It might do for undergraduate courses for non-scientists — in American terms, for fuzzies, non-science majors — but serious would-be scientists, the techies, haven't even got time for all the real stuff. In 1965, in his textbook for undergraduates *Molecular Biology of the Gene*, James Watson presented the science as though it were a mediaeval cathedral, the edifice inspiring, its artisans and architects anonymous (with exceptions including himself and Francis Crick). The book went through multiple editions.

Yet the alternative has occasionally been attempted. In 1971, Gunther Stent published *Molecular Genetics: An Introductory Narrative*. "The evolutionary origin and essentially pedagogic purpose of this book are reflected in the narrative presentation of the material in the historical sequence in which it actually came to be known (and in the occasional burdening of the reader with long-abandoned theories)," Stent wrote. "I happen to believe that an understanding of molecular genetics can best be taught in an organic (rather than logical) manner." Helped, to be sure, by muscular writing and Stent's immense acquaintanceship with the artisans and architects, the book stands as a vindication of the historical approach. It sold well enough to warrant a second edition before the subject grew too cumbersome.

Now we have Harrison Echols' *Operators*

and *Promoters*. Its subtitle claims that it's a history; its very format raises doubts. It weighs in at a kilogram, with heavy paper, margins fully as wide as the column of type, and a pricey illustration programme. Yet the book asserts that historical claim repeatedly, in a foreword by Arthur Kornberg, an afterword by Tom Cech and a long, laudatory preface by Carol Gross, Echols' widow and editor.

At the University of California, Berkeley, Echols was at one of the nodal points in the worldwide network of the research he has written about. In 1986, he learned he had cancer. For the next six years, Gross writes, his overriding aim was to complete "a personal account of the development of the biological paradigms that we now take for granted". But when he died, in 1992, he left a mass of materials, unfinished and unpolished. Echols' widow, herself a molecular biologist at the University of California, San Francisco, took it on. Echols wrote, Gross says, from his own scientific work, from his encounters and interviews with other biologists, and "with the eye of a physicist, the knowledge of a biologist, and the soul of an artist".

With respect, the history of science demands more than the careful ascription of results to individuals, leavened with the

## New in paperback

**Planetary Dreams: The Quest to Discover Life Beyond Earth**

by Robert Shapiro  
Wiley, \$16.95, £12.50

**Survival Strategies**

by Raghavendra Gadagkar  
Harvard University Press, \$17.95, £12.50

**Travels to the Nanoworld: Miniature Machinery in Nature and Technology**

by Michael Gross  
*Perseus*, \$16, £11.99  
"The style of writing is hard to place ... although Gross's humour and enthusiasm are considerable compensation ... This book provides captivating snapshots, but sometimes lacks the narrative of a good movie, or book for that matter." Philip Ball, *Nature* **402**, 119–120 (1999)

**The Scalpel and the Butterfly: The Conflict between Animal Research and Animal Protection**

by Deborah Rudacille  
University of California Press, £12.95, \$17.95

**Zero: The Biography of a Dangerous Idea**

by Charles Seife  
Souvenir Press, £9.99, \$13



occasional personal anecdote or quotation. It requires detachment, selectivity and an explicit understanding not just of the building of the science and its subsets, stone by stone, but of the styles of the different institutions where the science was done, and of the evolution of the structure of the field, both intellectually and as that network of interactions.

The history of molecular biology is marked above all by a radical shift in the nature of the central problems and by a concomitant transformation of the research networks. By 1970, the founding fathers had erected an elegant, overarching outline of the nature of the genetic material and the relationships between genes and their products, between DNA, RNA and proteins, with prototypical forms of genetic regulation. So the chief task became, in Crick's phrase at the time, to do molecular biology all over again for eukaryotes. This has meant turning at last to what has for centuries been biology's great intractable, the working out of what used to be called embryology — development and differentiation.

The other aspect of the transformation has been, of course, the splintering of the research into multiple lines and the high-exponential proliferation of the numbers of workers and research centres. Yet all those lines and most of those workers must be concerned primarily with the control of the expression of genes — not just in elegant outline but in full and stupefyingly complex detail. Stent published his narrative textbook on the cusp of the transformation. His future was Echols' and Gross's past. The bulk of their book treats a selection of these mechanisms of control. This treatment is right in principle, yet the selection is skewed by its tight dependence on Echols' own research background. The result, if it qualifies as history at all, is internalist to an extreme, presenting the science narrowly and technically — just what textbooks do.

We are indeed told who did what. But operators and promoters? That could be piquant. Science at its creative best is as dependent as, say, sculpture or theatre or pop music on the styles of individuals and groups. The double meaning of the title was deliberate, yet Echols and Gross reveal little of the individuality of scientists. They tell us virtually nothing about where they worked, or how they interacted with colleagues and competitors. Nor do they address its societal context, how the science described in the book has affected and been affected by its rise to its present dominance of the headlines and of public policy. Instead, those wide margins are decorated by pencil portraits of molecular biologists — 76 of them, including three women. I know 43 of them: only about six are recognizable. But they're smiling. ■

Horace Freeland Judson is at the Center for History of Recent Science, George Washington University, Washington DC 20052, USA.

## Science in culture

### Bee string theories

*Humble Boy*, a play by Charlotte Jones.

Sara Abdulla

"I have been doubly unlucky in life: to marry a biologist and give birth to a physicist." So says Flora Humble, the linchpin of Charlotte Jones's gentle new comedy. Flora's astronomer son — the 'Humble Boy' of the title — is home for her entomologist husband's funeral. For one long, hot summer, mother and son grope through grief towards an understanding of their own and the dead man's life, motivations and legacy.

*Humble Boy* is an entertaining celebration-cum-send-up of the personal and professional quest for immortality through science, love and children. More a laid-back buffet in the manner of Yasmina Reza's recent *Lifex3* than an earnest essay à la Michael Frayn's *Copenhagen*, it is a likeable domestic drama. It contrasts the personal satisfactions of amateur research (unbeknownst to his family, the late James Humble discovered a new strain of bee) with the gallery-playing antics of the academic life (Felix Humble's febrile pursuit of Cambridge kudos). The play pits the palpable beauty of the natural sciences (the garden setting is lovingly Latin-labelled by James's ghost) against the poetry of mathematical astrophysics (Felix transcends his stutter to enthuse on the equations describing the Universe's coiled dimensions).

Magpie-like, Jones picks shiny bits of grade-school science in a shameless but irresistible attempt to graft on gravitas. Her characters inter-

relate like the insects in James Humble's centre-stage hive. The indomitable Flora is the queen bee — life-giving and languid. Her husband, lover and even her son are drones bent to her bidding. Her childless female friend — her worker.

Similarly, the play's emotional journey skims off superstring theory's easy-to-digest cream. Professionally, Felix Humble is searching for a 'Theory of Everything' that will reconcile general relativity with quantum mechanics. This is echoed by his personal struggle to understand how the big events in life — births, deaths, marriages — are informed by the seemingly inconsequential stuff. It may sound trite but, thanks to Jones's facility with dialogue, it works.

Less successful is her recourse to tired tropes, such as the socially inept researcher evading his emotional shortcomings through a life of the mind. That her women find creative fulfilment and make their mark solely by having babies also jars. Jones redeems herself by painting parenthood as "full of Eureka moments", thanks to the insatiable — scientist-like — curiosity of children. Profound? Not really. Pleasurable? Definitely. ■

Sara Abdulla is editor of the *Nature News Service*.

*Humble Boy* is at London's Cottesloe Theatre.

Visualizations: *The Nature Book of Art and Science*, a collection of essays edited by Martin Kemp, is published by Oxford University Press (£20) and the University of California Press (\$35).

Simon Russell Beale as Charlotte Jones's Felix Humble.



# Good news is no news

How can scientists use the media to give their side of the story to the public?

John Emsley

**T**alk to a chemist about the media and he or she will complain that their science gets a raw deal. And so it is, and has been for 20 years or more. A typical example in the United Kingdom is the *Daily Mail*, whose front page, one day in March 1996, shrieked “CHEMICAL TRAIN SMASH HORROR!” In fact, the chemical in question was carbon dioxide: a tank of it had ruptured and was “spewing gas all over the place”, creating “a 3-ft deep mist”. That was all—but it merited an alarmist headline. And there’s the rub: chemical rubbish gets publicized while real chemistry goes unreported.

The subeditor who devised the above headline was catching the readers’ attention with a carefully baited emotional hook, suggesting pollution and destruction. Pressure groups who seek to capture public notice are skilled in providing such hooks, often to publicize opinions that infuriate scientists with their misinformation and disinformation. Chemists suffer more than most. They are blamed by environmentalists for causing pollution, by alternative-health gurus for causing cancer, and by organic-food producers for contaminating crops. All these groups put out regular press releases, with excellent hooks.

Rational argument alone will not carry a message to the general public; it has to travel on the back of emotion. People will understand that medical or pharmacological advances offer a new treatment for a particular disease, but the information will be quickly forgotten unless the news has a visceral component that registers feelings such as fear for oneself or sympathy for others.

There can be various degrees of emotional involvement in a news item, ranging from the personal, through the societal to the global. A common ploy of the first type is to reveal a threat to a group we are bound to be sympathetic towards, such as babies, breast-feeding mothers or young children. Threats to pregnant women or to male fertility are

likely to hook the young, while risks of heart disease and cancer hook those who are older. In the societal category the targets are broader and consist of threats to food, wealth and health, while for global hooks the emphasis is on deprived people in developing countries, or on wildlife, or on the planet as a whole.

There is no reason why stories from scientists should not use the same emotional hooks to get their message across. For example, advances in agrochemical fertilizers and pesticides can be justified as saving wildlife because they allow more human food to be produced on less land, thereby reducing the destruction of natural habitats. A new drug may have economic benefits, for example by reducing the need for a stay in hospital, but it would be better to say that it will transform the health of the afflicted and concentrate on their currently blighted lives.

All this may sound a bit contrived, and you may think that if people don’t want to read about sound science, then it’s their loss and that’s the end of the matter. But the mainstream sciences will never regain the standing they once had if we continue to communicate badly. Nor is it a hopeless task. Even chemistry has had its news successes, the most notable being Viagra, which captured wide media attention in the late 1990s. There, of course, the hook was the most powerful of all: sex.

Yet I am sure that many molecular scientists will remain unconvinced by the need to be pro-active in communicating. It might help them if I draw an analogy between communication and oxidation. No matter how strong the oxidizing agent (the information) it cannot attract electrons (attention) from some atoms or molecules. What may be needed is something that starts the process by tempting an electron into an outer orbital (the hook).

Even so, putting across good news has an added energy barrier to surmount: only bad news is good news as far as the popular press is concerned. Those who demonize the products of science have become skilled in the art of writing press releases full of doom and gloom. “Babies at risk from PVC toys” is a typical example which ran in the media last Christmas, and used a well-trying hook. The publicists of organizations issuing such warnings can generally come up with skilfully contrived phrases that even enter everyday speech, such as ‘gender-benders’, ‘Frankenstein foods’, ‘cancer on tap’ and ‘artificial fertilizers’—all highly emotive, and all misleading.

If we want sound science to make the



headlines in a positive way, we need to think of how we package it. OK, you’ve done some brilliant research and now you want your company, university or research institute to issue a press release about it. Of course, you can simply describe what you have done, say who funded the research and what its aim was, add a brief description of yourself and give a contact number, and wait for the journalists to ring you. And wait, and wait, and wait... what your story lacks is a good hook.

So what are the good hooks? The top three are sex, money and health. Link your story to one of those and you’re away. The newspaper subeditors who devise headlines will have no difficulty in catching the reader’s attention with SEX, CASH and CURE. Other hooks can be almost as effective. Have your research findings explained the unexplained? Or are they contrary to expectations? These make intriguing hooks with headline words such as CLUE and SHOCK. A local connection, a research first, possible job creation and even national pride are also useful areas to emphasize.

Failing any of these, you should try to suggest that your work will provoke controversy. In the United Kingdom, headline writers are wedded to the hook words of confrontation, such as FURY, FEAR, STORM, ROW and LOOMS. For example, to create a news item about this article, I would find someone who disagrees with something I have said, and then head a press release about it with the phrase: NEWS SPIN ROW LOOMS. Shock horror—it can be done. ■

*John Emsley is science writer in residence in the Department of Chemistry, Cambridge University, Lensfield Road, Cambridge CB2 1EW, UK, and author of Nature’s Building Blocks: An A–Z Guide to the Elements (Oxford University Press).*

DAVID NEWTON

If we want sound science to make the headlines in a positive way, we need to think of how we package it.



# Fidelity and infidelity

Miroslav Radman

**T**he apparent perfection of organisms and the accuracy of biological processes are still used in religious explanations of the origin of life. However, in real life it is survival, not fidelity, that is the ultimate virtue.

Because adaptability involves exploration of genetic possibilities to fit ecological niches, molecular infidelity and repetition are more likely to succeed than a precise, non-repetitive process. Only a tiny fraction of antibodies produced will ever be useful; the rest can be considered as mistakes. At least half of all human embryos fail during development. During chromosome segregation from a mother cell into two daughters, the polymerizing fibres (microtubules) do not know the exact location of the chromosomal target (the centromere) — they shoot and miss until one hits. A precise, single shot would often miss a target of uncertain position, whereas successive, imprecise firing will eventually lead to a hit. Selection at the level of molecules, cells and organisms may give the impression of designed perfection, but life's structures do not emerge by fully deterministic design.

What about the fidelity of enzymatic reactions? In the very precise process of DNA

replication, accuracy is achieved by using a 'proofreading' system to remove erroneously inserted nucleotides, and then by quality-checking the synthesized DNA using a mismatch-repair system that removes virtually all remaining mistakes. It would take too long to get it exactly right in the first place. DNA replication is efficient and therefore relatively imprecise, leaving mistakes to error-correction enzymes which are themselves efficient because their substrates are specific mistakes made by other enzymes.

John Hopfield and Jacques Ninio introduced the concept of 'kinetic' proofreading to explain the fidelity of molecular processes in which mistakes cannot be identified once they have been made. By delaying formation of the product, erroneous substrates are less likely to be transformed into it. Such mechanisms are seen in protein biosynthesis and may also be involved in presentation of antigens to T-cell receptors. In these cases, high fidelity is gained at the expense of efficiency.

Does nature attempt to achieve perfect fidelity using repair enzymes? Apparently not, because some bacterial mutants have increased fidelity of DNA or protein biosynthesis — an example is so-called StrR mutants, which are resistant to the antibiotic streptomycin. Streptomycin increases the error rate in protein biosynthesis; in its absence, StrR mutants have high fidelity, and grow more slowly than non-mutants. Because fidelity costs, it is optimized rather than maximized. Of course, error rates that are too high can cause death — bacteria and haploid yeast die from a genetic-error catastrophe when their mutation rate is increased 10,000-fold.

There is ample evidence of trade-offs between accuracy and efficacy. In protein synthesis, there are about 3 erroneous amino acids for every 10,000 correct ones; in transcription, about 1 nucleotide per 100,000 is wrong; and in DNA replication, about 1 nucleotide is wrong in every 10 billion. Proteins are functional molecules that are used up, oxidized, broken down and replenished by *de novo* synthesis. One-third of all proteins synthesized by normal human cells are immediately degraded by proteasomes because they have recognizable folding mistakes. Up to half of all messenger RNAs are erroneous and are therefore broken down. These clean-up systems probably contribute to the functional longevity of non-dividing cells such as neurons and cardiac muscle.

There is a huge investment in the fidelity of DNA replication, because genes are life's database. But when survival is threatened, even DNA fidelity is relaxed. For instance, when bacterial adaptation is limited by the available genetic diversity, genetically unstable

## Molecular evolution

*Errors and infidelity, even wastefulness, can cause individual failure, but they are also a source of innovation and robustness, ensuring the perpetuation of life.*

populations adapt, at least in the short term, more rapidly than their stable counterparts.

Two classes of genes accelerate genetic variation by enhancing mutation and/or recombination in bacterial populations: stress-inducible wild-type genes and genes whose functional loss increases genetic variability; the former acts on individual cells and the latter on populations. Furthermore, genes encoding proteins under persistent selective pressure, such as a host's immunoglobulin genes or parasite target proteins/genes, show increased mutation rates. Thus, the fidelity of individual genes can be optimized.

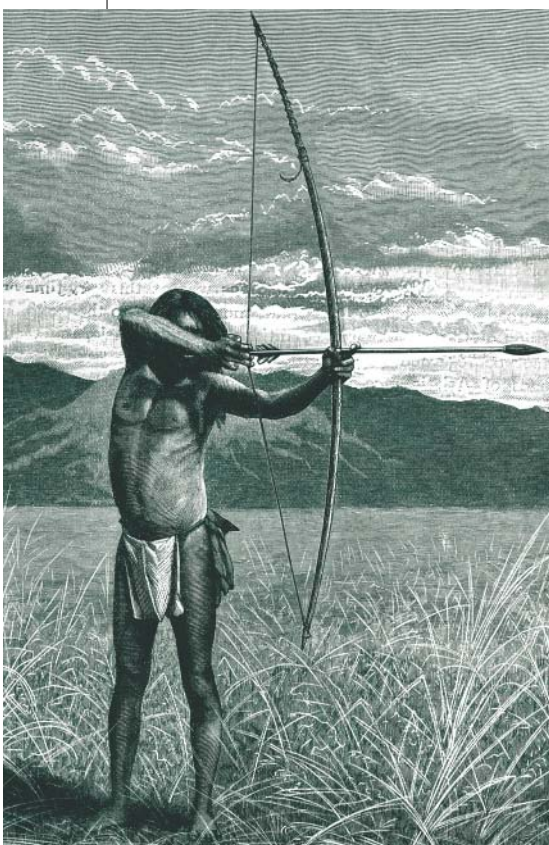
There is also a down-to-earth reason for DNA infidelity: DNA replication itself! A top-fidelity DNA-copying machine, such as the DNA-replication complex, can only copy chemically perfect DNA. But cellular DNA is not perfect, because of spontaneous oxidation, hydrolysis, alkylation, strand breaks and so on. As many as 300,000 such lesions may occur every day in every cell of a human body, most of which are dealt with by special repair systems. Each unrepaired lesion is a potentially lethal event that can stop the regular DNA-replication machinery. However, SOS, a class of unusual DNA polymerases, allows DNA replication to proceed despite the lesions.

Errors and infidelity, even wastefulness, can cause individual failure, but also provide innovation and robustness, ensuring the perpetuation of life. Nature does not exhaust itself for the sake of fidelity and perfectionism. Rather, errors are made, often repaired or discarded, but always tested as the source of blind innovation during the continuous adaptation to unpredictable environmental changes and challenges. ■

Miroslav Radman is at the Laboratory of Evolutionary and Medical Molecular Genetics, INSERM E9916, and in the Department of Cell Biology, Université de Paris-V, 75730 Paris Cedex 15, France.

### FURTHER READING

Caporale, L. H. (ed.) *Molecular Strategies in Biological Evolution* (New York Acad. Sci., 1999).  
Keller, E. F. *The Century of the Gene* (Harvard Univ. Press, Cambridge, Massachusetts, 2000).  
Freidberg, E. C. *Correcting the Blueprint of Life* (Cold Spring Harbor Lab. Press, Cold Spring Harbor, New York, 1997).



**Moving target: roughly aimed shots lead to a hit.**

# Particles driven to diffraction

Philip H. Bucksbaum

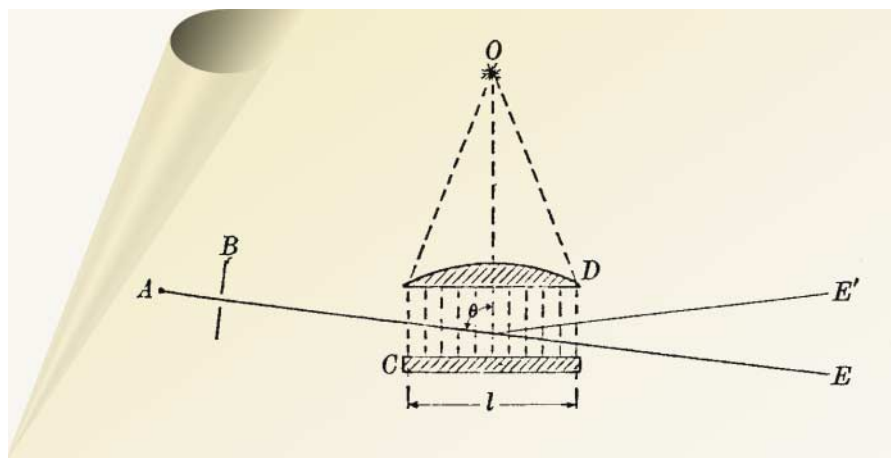
Almost 70 years after it was first proposed, an experiment shows that electrons can be diffracted by light waves. This result highlights the interchangeable roles of matter and light.

Wave-particle duality is the concept that all particles can behave as waves, and vice versa. This intellectually challenging notion, which is a fundamental prediction of quantum theory, has been tested in a new way by Herman Batelaan and co-workers, in an experiment reported on page 142 of this issue<sup>1</sup>.

The debate over the particle versus wave character of light is far older than quantum theory. Newton was an early and active advocate for the corpuscular nature of light. But it was in the first decades of the twentieth century that quantum mechanics brought this discussion to a new plane by including matter as just another form of energy subject to the wave-particle dichotomy.

What does it mean to say that matter behaves like a wave? We know waves have ripples but, because atoms and electrons are so small, if they are waves then their ripples must be tiny. The quantum ripples of an electron in an atom are typically less than an ångström, or one ten-billionth of a metre, in size. But we don't need to see ripples to detect waves. The accepted evidence for wave-like behaviour is the phenomenon of diffraction.

Diffraction is easy to demonstrate for light. The



**Figure 1 Making light of the matter.** Below left, diffraction of light forms a rainbow pattern on the surface of a compact disk. Above, a drawing from Kapitza and Dirac's 1933 paper<sup>2</sup> that describes a proposed method for the diffraction of electrons (from the path AE to AE') from an optical standing wave formed by a light source, O, a collimating lens, D, and a mirror, C. Batelaan and co-workers<sup>1</sup> use a similar geometry in their experiment.

rainbow pattern of colours that you see when you look at the surface of a compact disk is caused by light waves diffracting from the regularly spaced bands of shiny material that make up the tracks. This effect can be seen because the wavelength of light, although small, is large enough to be comparable to the spaces between adjacent tracks.

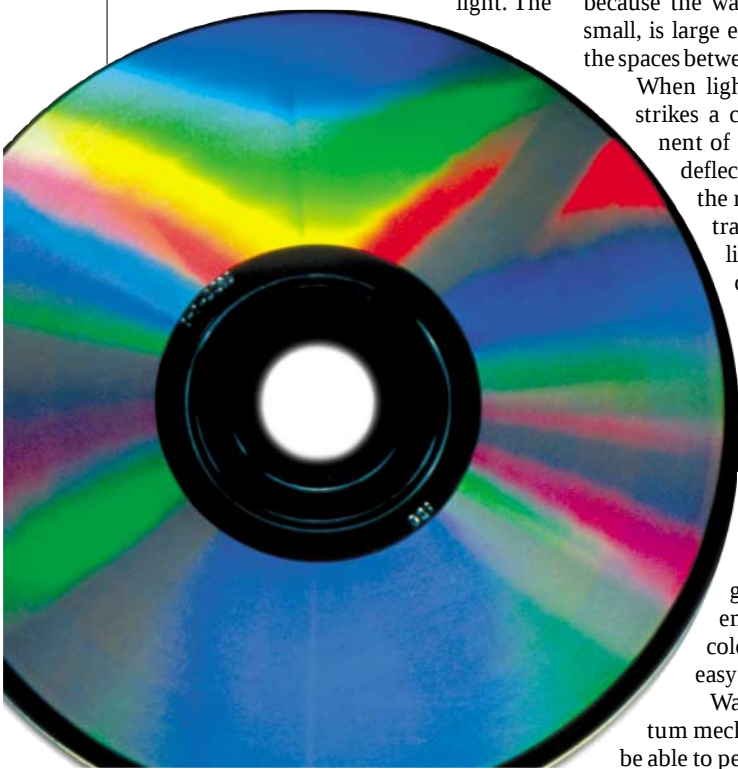
When light from a lamp or the sun strikes a compact disk, each component of colour in the 'white' light is deflected in a direction dictated by the ratio of its wavelength to the track spacing. Specifically, for light with wavelength  $\lambda$  incident at  $90^\circ$  on a grating with track spacing  $d$ , diffraction occurs at an angle given by  $\sin\theta = \lambda/d, 2\lambda/d, 3\lambda/d$  and so on. Light waves scattering from all of the tracks add coherently only at these special angles. The wavelengths of visible light are tiny (just 400–700 nanometres) but, if the grating spacing  $d$  is small enough, the separation of the colours (due to the angle  $\theta$ ) is easy to detect.

Wave-particle duality in quantum mechanics means that we should be able to perform the same observation

as described in the previous paragraph, even when the light waves are replaced by particles and the material grating by light. Think about how to make that compact disk out of a light beam for a moment. Don't panic if you haven't come up with a solution; the Batelaan group<sup>1</sup> has done it for you.

Batelaan and colleagues used a method originally proposed by two brilliant physicists, Paul Dirac and P. L. Kapitza, in a classic paper<sup>2</sup> written in 1933. Dirac and Kapitza each won Nobel prizes later, but not for this work. This is the only paper that they wrote together, and it seems to be an isolated curiosity. It was not written to resolve the wave-particle debate, because by the early 1930s this had been decided by numerous experiments in favour of... well, both particles and waves, as quantum theory predicts. Nonetheless they wrote that the diffraction of electrons by light would be a very interesting experiment.

The figure in the Kapitza–Dirac paper reveals the trick for creating a regular lattice of optical radiation (Fig. 1). Kapitza and Dirac reasoned that an optical standing wave would have the correct properties. A standing wave is just wave-like motion that oscillates but doesn't travel, such as the oscillations of a vibrating violin string. An optical standing wave has an oscillating electric field made by two counterpropagating and overlapping light beams. Kapitza and Dirac





suggested that a standing wave of light could be constructed from radiation produced by mercury atoms in an arc lamp, which fluoresce in intense and sharp wavelength bands. If you use a compact disk to diffract light from a fluorescent lamp, you should be able to see the diffraction lines they were thinking about, because most fluorescent lamps produce their light from mercury atoms.

An electron beam in which the electrons follow parallel paths (are collimated) and have similar velocities would diffract from the standing wave at angles given by the same formulae that describe the diffraction of light from a grating — that is, at an angle determined by the ratio of the electron's wavelength to the period of the standing wave. (The period of the light grating is half the optical wavelength, because there are two intensity peaks per cycle in a standing wave.) The electron's quantum mechanical de Broglie wavelength, according to quantum theory, is inversely proportional to its momentum  $p$ :  $\lambda_{\text{deBroglie}} = h/p$ , where  $h$  is Planck's constant. So the angle of deviation for an electron beam incident at  $90^\circ$  should be integer multiples of  $2h/(\lambda p)$ , where  $\lambda$  is the wavelength of the light. This is one-hundredth of a degree or so for a grating of green light and electrons with 380 electron volts of energy, as in the Batelaan experiment.

But there is a problem: the force exerted by optical radiation on free electrons is incredibly weak. In other words, returning to our original experiment with the compact disk, it is as though the disk were nearly invisible because it was made of something with nearly the same optical properties as the air around it. In that case the light would pass right through it, and never diffract. This is why the Kapitza–Dirac thought experiment remained untested for many decades.

The situation improved following the invention of the laser. Continuous semiconductor diode lasers like the ones that read compact disks are still not powerful enough to demonstrate the Kapitza–Dirac effect for free electrons, but they can be used to diffract beams of atoms if their wavelength is close to an atomic transition line (like the spectral lines you get from fluorescent mercury atoms). Diffraction of matter waves by optical standing waves was therefore first demonstrated using a beam of neutral atoms passing through the optical standing wave of a continuous laser<sup>3</sup>. In this experiment, Gould, Ruff and Pritchard confirmed the Kapitza–Dirac formula, and in so doing helped to stimulate interest in the kind of particle–wave physics known today as atom optics. Techniques of atom optics have led to atomic Bose–Einstein condensates, atom lasers and developments in direct-write atom lithography.

The scattering force on free electrons by laser light is more than a billion times smaller than the carefully tuned laser–atom force,

so the light has to be much more intense to have any effect. Continuous lasers simply cannot be made strong enough, but pulsed lasers can bridge this gap easily. The effect of light on free electrons was first observed<sup>4</sup> using large pulsed lasers in the 1960s, and it was studied in detail in the 1980s, partly by some experiments using standing waves<sup>5</sup>. These experiments did not use sufficiently collimated electrons to see the individual peaks at different angles that are the hallmark of electron diffraction from a standing wave. The full experimental test of this idea has had to wait until now, 40 years after the invention of the laser, and nearly 70 years since the Kapitza–Dirac paper.

Batelaan and colleagues' experiment<sup>1</sup> is well executed and the series of electron peaks at different scattering angles seen in Fig. 2 on page 143 is in beautiful agreement with the Kapitza–Dirac theory. A larger question, though, is where this advance leads us in physics. Batelaan's group proposes using it as a spectroscopic tool, or using the multiple

peaks to build electron interferometers. These ideas are worth pursuing. They belong to a new field of physical research devoted to manipulating quantum phenomena using the exquisite control we now have over laser fields. Quantum computing, slow light, atom lasers and similar subjects that have appeared in these pages in the recent past belong to this new field of quantum control. The Batelaan experiment helps to tie these new advances to the foundations of quantum theory. ■

Philip H. Bucksbaum is at the FOCUS Center, Department of Physics, University of Michigan, Ann Arbor, Michigan 48109–1120, USA.

e-mail: phb@umich.edu

- Freimund, D. L., Aflatooni, K. & Batelaan, H. *Nature* **413**, 142–143 (2001).
- Kapitza, P. L. & Dirac, P. A. M. *Proc. Camb. Philos. Soc.* **29**, 297–300 (1933).
- Gould, P. L., Ruff, G. E. & Pritchard, D. E. *Phys. Rev. Lett.* **56**, 827–830 (1986).
- Bartell, L. S., Roskos, R. R. & Thompson, H. B. *Phys. Rev.* **166**, 1494–1504 (1968).
- Bucksbaum, P. H., Schumacher, D. W. & Bashkansky, M. *Phys. Rev. Lett.* **61**, 1182–1185 (1988).

## Physiology

# Nitric oxide and respiration

Stuart A. Lipton

The theory that haemoglobin evolved to carry oxygen around the body may need a rethink in light of another way in which molecules related to nitric oxide, released from haemoglobin, help the brain control respiration.

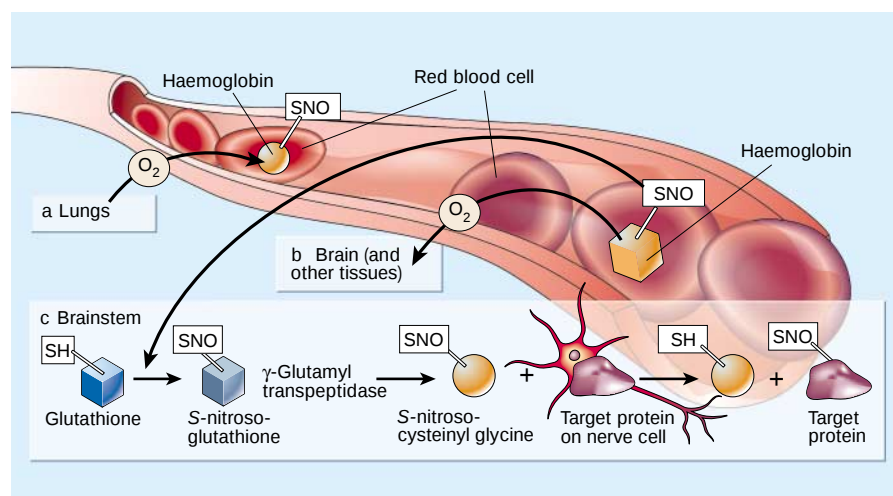
Our ability to breathe, and to modify our breathing according to the amount of oxygen in the air and the demands of our bodies, is essential for survival. Failure to breathe more often when oxygen levels are low can contribute to respiratory distress in newborn mammals and to sleep apnoea — temporary inability to breathe — in adults. How is the increase in breathing in the face of a shortage of oxygen controlled? As Lipton (no relation!) and colleagues<sup>1</sup> tell us on page 171 of this issue, the answer seems to be not by the mere lack of oxygen, but rather by molecules related to a different gas, nitric oxide (NO), which affect respiratory centres at the base of the brain.

The molecules in question are S-nitrosothiols (SNOs) — complexes of NO bound to a thiol (sulphydryl) group in the amino acid cysteine. This is not the first time that these molecules have been found to be involved in respiration. They also have a crucial role in matching ventilation to perfusion in the lungs (that is, in matching the diameter of the airways to the diameter of the lung blood vessels)<sup>2</sup>. This ensures that haemoglobin, the blood's oxygen-carrying molecule, is fully loaded with its cargo.

Moreover, SNOs are involved in controlling the supply of oxygenated blood to tissues. In the lungs, SNOs and oxygen are

loaded onto haemoglobin together; SNOs are released from haemoglobin when it is deoxygenated, dilating the small blood vessels that deliver oxygen directly to tissues<sup>3,4</sup>. Remarkably, then, SNOs seem to control all three parts of the respiratory cycle: the oxygenation of haemoglobin in lungs, the delivery of oxygen to tissues, and the control of breathing by the brain.

In an elegant series of experiments, ranging from physiological to chemical analyses, Lipton *et al.*<sup>1</sup> show that increased breathing ('minute ventilation') is controlled in part by SNOs acting in an area called the nucleus tractus solitarius in the brainstem. The authors also tracked down the exact molecule — a metabolite of S-nitrosoglutathione — that probably exerts this ventilatory effect. The metabolite is S-nitrosocysteinyl glycine, which is produced in neuronal tissue through cleavage of S-nitrosoglutathione by the enzyme  $\gamma$ -glutamyl transpeptidase. This requirement for S-nitrosoglutathione and  $\gamma$ -glutamyl transpeptidase may distinguish the effects of SNOs in the brainstem from their oxygen-regulated effects on blood vessels — the latter effects are not reproduced by application of S-nitrosoglutathione<sup>3,4</sup>. In future, it might be possible to take advantage of this distinction to inhibit or augment the effects of specific SNOs on particular aspects of respiration.



**Figure 1** How S-nitrosothiols (SNOs) work on the brainstem to control breathing. **a**, Haemoglobin is oxygenated in the lungs; one of its cysteine residues is then modified with SNO. **b**, Oxygen levels in the brain (and other peripheral tissues) are naturally low. Haemoglobin releases its oxygen and changes its shape in a way that ultimately promotes transfer of its NO group to another cysteine residue, in this case on glutathione (**c**), by transnitrosylation. S-nitroso-glutathione is cleaved by the enzyme  $\gamma$ -glutamyl transpeptidase in neuronal tissue to form S-nitrosocysteinyl glycine. It is also possible that S-nitrosocysteinyl glycine is then cleaved by a peptidase to produce S-nitroso-L-cysteine. One of these two molecules is then proposed to transnitrosylate a cysteine residue in a target protein in the nucleus tractus solitarius region of the brainstem, regulating the central neuronal circuitry that controls breathing.

In summary, the work of Lipton *et al.*<sup>1</sup> will make it into the physiology textbooks with four key observations. First, a blood-derived factor can control minute ventilation by acting on neurons in the nucleus tractus solitarius. Second, specific SNOs are involved. Third, enzymes ( $\gamma$ -glutamyl transpeptidase and possibly others) are needed to transfer the SNO signal from blood to brainstem. Finally, haemoglobin releases SNOs not only to produce local blood-vessel dilation — to match the oxygen requirements of different tissues to blood flow — but also to control areas in the central nervous system that are involved in breathing.

Given these results, we may legitimately question whether haemoglobin first evolved to carry oxygen or to ferry NO to key locations in the body. It has been argued that haemoglobin's original task was to enable reactions involving NO and SNOs to occur<sup>5</sup>, and that its ability to carry oxygen came later; other work supports this view<sup>6</sup>. In fact, many responses to low oxygen levels (hypoxia) can be attributed to the SNO-generating transfer of NO groups from haemoglobin; these responses include those discussed above<sup>1,3,4</sup> as well as modulation of gene expression, which is mediated by increased activity of the transcription factor HIF-1 (ref. 7). Expression of these genes may also contribute, directly or indirectly, to hypoxic preconditioning — a mechanism whereby short exposures to hypoxia or ischaemia (lack of blood flow) can protect the heart and brain from damage resulting from more prolonged hypoxia<sup>8</sup>.

The idea that the respiratory response

to hypoxia is controlled by SNOs is radical, not only because NO was once thought to be inactivated by haemoglobin, but also because oxygen has always been viewed as central to cellular respiration and energy production. Nonetheless, it seems clear that NO-related molecules produce a web of control over the level, delivery and metabolism of oxygen. Oxygen works in concert with, and under the control of, SNOs.

How exactly do SNOs reach the brainstem from the blood? It seems likely that the mechanism involves transnitrosylation — the transfer of SNOs from a key cysteine residue of one peptide to another. A series of such reactions is apparently required to transfer the NO group from haemoglobin, first to the red-blood-cell membrane (probably through anion-exchange protein-1)<sup>4</sup>, then onto glutathione in the blood and into endothelial cells (which line the walls of blood vessels)<sup>9</sup>, and finally generating S-nitrosoglutathione in the brainstem (Fig. 1).

The process then requires  $\gamma$ -glutamyl transpeptidase to produce S-nitrosocysteinyl glycine, which might likewise be processed to form S-nitroso-L-cysteine, a molecule that has long been known to act in a stereospecific manner to reduce heart rate and blood pressure when injected into the nucleus tractus solitarius<sup>10,11</sup>. One interpretation of this stereospecificity is that there is a receptor for L- (but not D-) SNOs on brainstem neurons. Another is that a transporter molecule is involved in importing S-nitroso-L-cysteine or S-nitrosocysteinyl glycine into neurons. But it is equally



#### 100 YEARS AGO

It has long been recognised that a change of surroundings may profoundly influence the reproductive system, in some cases increasing the fertility, in others leading to complete sterility... An Arab-Kattiawar pony which arrived during April from India proved during the first three months quite sterile, owing, I believe, to loss of vigour on the part of the germ-cells, their vitality being only about one-tenth that of a home-bred hackney pony. But the fertility is apparently impaired by even comparatively slight changes of environment. Lions which breed freely in Dublin seem to be sterile in London, and I heard recently that when bulls are changed from one district to another in the north of Ireland complete sterility is sometimes the result... No one doubts that the bodily vigour is liable to be impaired by fevers and other diseases, by changes in the habitat, unsuitable food, rapid and unseasonable changes of temperature, and the like; hence it will not be surprising if further investigations prove that changes in the soma, beneficial as well as injurious, are reflected in the germ-cells, and thus indirectly induce variation.

From *Nature* 12 September 1901.

#### 50 YEARS AGO

Early last year Sir Stanley Unwin reviewed in a pamphlet entitled "How Governments Treat Books", some of the obstacles which taxation policy offers today to the free flow of books from one country to another... Sir Stanley in this pamphlet, rightly emphasizing the value of books for the transmission not merely of knowledge but also of thought, pointed out that post-war governments are treating with scant respect this unique and priceless instrument for giving continuity and permanence to the free expression of thought. In many otherwise civilized countries, there is an increasing tendency to treat books as if they were merely an ordinary commodity of commerce, without cultural value or importance. The fundamental principle that the one thing no country can afford to hamper, limit or tax is knowledge no longer receives the respect that it enjoyed a century ago; and in the climate of opinion that is engendered by security measures, once they escape beyond the limits that the national interest requires in such matters as atomic energy, the progress of knowledge is bound to slow down.

From *Nature* 15 September 1951.



plausible that transnitrosylation is stereoselective, as a consequence of the quaternary structure of the peptides involved, and that L-SNOs are simply more effective for transnitrosylation of the ultimate molecular target in the brainstem.

What is this final target? Several receptors in the central nervous system are S-nitrosylated as a means of controlling their activity. Chief among them is the *N*-methyl-D-aspartate subtype of receptor for the excitatory neurotransmitter glutamate<sup>12,13</sup>; this receptor is enriched in the nucleus tractus solitarius. But other targets could be involved. Cysteine residues must be present in certain motifs (specific sequences of amino acids) to accept SNOs and regulate protein function<sup>14</sup>. This motif is found in several thousand proteins<sup>14</sup>, and the importance of S-nitrosylation in the functions of these targets is just beginning to be determined. SNOs might also act outside the brain on other neurons, such as those in the carotid body, that project to the nucleus tractus solitarius. So far, the target proteins for SNOs in the brainstem have not been identified.

For now, however, we can take solace in

knowing that there is a grand scheme for coordinating perhaps the most fundamental of physiological processes — the response to oxygen deprivation — and the effectors are SNOs.

Stuart A. Lipton is at the Center for Neuroscience and Aging, The Burnham Institute, La Jolla, California 92037, USA.

e-mail: [slipton@burnham.org](mailto:slipton@burnham.org)

1. Lipton, A. J. *et al.* *Nature* **413**, 171–174 (2001).
2. Moya, M. P. *Proc. Natl. Acad. Sci. USA* **98**, 5792–5797 (2001).
3. Jia, L., Bonaventura, C., Bonaventura, J. & Stamler, J. S. *Nature* **380**, 221–226 (1996).
4. Pawloski, J. R., Hess, D. T. & Stamler, J. S. *Nature* **409**, 622–626 (2001).
5. Minning, D. M. *et al.* *Nature* **401**, 497–502 (1999).
6. Hausladen, A., Gow, A. J. & Stamler, J. S. *Proc. Natl. Acad. Sci. USA* **98**, 10108–10112 (2001).
7. Palmer, L. A., Gaston, B. & Johns, R. A. *Mol. Pharmacol.* **58**, 1197–1203 (2000).
8. Digicaylioglu, M. & Lipton, S. A. *Nature* **412**, 641–647 (2001).
9. Ramachandran, N., Root, P., Jiang, X., Hogg, P. J. & Mutus, B. *Proc. Natl. Acad. Sci. USA* **98**, 9539–9544 (2001).
10. Ohta, H., Bates, J., Lewis, S. & Talman, W. *Brain Res.* **746**, 98–104 (1997).
11. Davisson, R. L., Travis, M. D., Bates, J. N., Johnson, A. K. & Lewis, S. J. *Am. J. Physiol.* **272**, H2361–H2368 (1997).
12. Lipton, S. A. *et al.* *Nature* **364**, 626–632 (1993).
13. Choi, Y.-B. *et al.* *Nature Neurosci.* **3**, 15–21 (2000).
14. Stamler, J. S., Toone, E. J., Lipton, S. A. & Sucher, N. J. *Neuron* **18**, 691–696 (1997).

## Evolutionary biology

# Sum of the arthropod parts

Mark Blaxter

Being an arthropod, with an external skeleton and jointed limbs, is a good thing in evolutionary terms. But the question of how the main groups of arthropods are related remains a subject of intense debate.

Arthropods are everywhere: there are millions of species, each often occurring in astronomical numbers, and they inhabit every niche in the biosphere. The evolution of the phylum Arthropoda is a long-standing area of study, but the interrelationships of the main subphyla are still the subject of vigorous debate<sup>1</sup>. That point is reinforced by the appearance in the past few months of four papers<sup>2–5</sup> on the subject, two of which — by Hwang *et al.*<sup>4</sup> and Giribet *et al.*<sup>5</sup> — appear on pages 154 and 157 of this issue. Each paper is internally consistent but disagreement remains, even on some of the central issues.

Extant arthropods are divided into five subphyla: Hexapoda (insects), Crustacea (crabs, shrimps, barnacles and woodlice), Myriapoda (millipedes and centipedes), Chelicerata (spiders, scorpions, mites and ticks) and Pycnogonida (sea spiders) (Fig. 1, overleaf). The success of arthropods has arisen from their jointed external skeletons and segmentally repeated appendages, which allow easy movement and manipulation of the surroundings in environments as different as the deepest seas and the

driest deserts. The arthropod subphyla are defined by the various types of body pattern, the diversity of which also offers rich ground for studying the evolution of development. Differentiation of segments and their appendages during development is achieved by interacting patterning systems mediated by genes such as the homeobox (Hox) genes. Major changes are often driven or revealed by changing Hox gene expression patterns<sup>6</sup>.

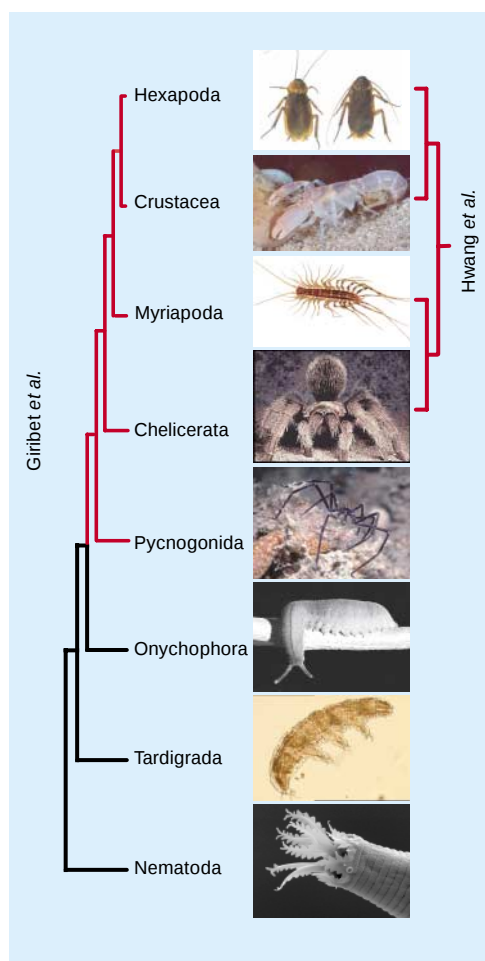
Interpreting the dynamics of morphological evolution requires a robust framework linking the subphyla. The splitting of the main arthropod lineages occurred over 540 million years ago<sup>7</sup>, around the beginning of the Cambrian period, and specialized arthropods were globally distributed a little later, suggesting that the group has a significant Precambrian history. However, the earliest multicellular fossils are found in rocks that are 560–590 million years old. If these fossils represent ancestral animals, this leaves only about 20–50 million years for the evolution and diversification of arthropods (and other modern multicellular animals). Molecular clocks, which use relative changes

in DNA and protein sequence to infer the time of separation of lineages, suggest that the diversification of arthropod subphyla occurred well before the beginning of the Cambrian, but with an earlier time for divergence (around 700 million years ago)<sup>8,9</sup>.

To resolve these closely spaced events, the data must retain a signal that reveals the evolutionary history despite the effects of over half-a-billion years of convergence (which generates apparent similarity from ancestral difference) and drift (which homogenizes and obscures overall divergence). Partial data sets can be highly misleading<sup>10</sup>, so large quantities of data from as many informative taxonomic groups as possible need to be included. By sampling more characters (more sequence from more genes, and more extensive morphological data sets) from more taxa, and with better computer modelling of the course of evolution, the 'shared, derived' characters of common ancestry may be revealed. Definition of shared, derived characters is essential: structures can be similar not because they have a shared ancestry, but because, as superficially similar solutions to life on a crowded planet, they have converged in appearance.

In addition, to resolve the internal structure of an evolutionary tree, the root of the tree must be defined with confidence. The position of the root will significantly affect how the evolution of various traits is modelled, and trees are best rooted using the closest related taxon, or outgroup. Molecular and morphological evolutionary analyses mostly agree that arthropods are part of a group of moulting animals, the Ecdysozoa<sup>11</sup>. The reality of the Ecdysozoa has just been reaffirmed by a herculean analysis of over 300 ribosomal RNA genes and 138 morphological characters by Peterson and Eernisse<sup>2</sup>. This work shows that the outgroups to use in tackling arthropod evolution are the Onychophora (velvet worms), Tardigrada (moss piglets or water bears) and Nematoda (roundworms). Using these outgroups, it may be possible to infer the structure and biology of a 'basal' arthropod with some confidence.

Recent work has shown that the Hexapoda and Crustacea are sister taxa ('Pancrustacea'), with hexapods sometimes 'nested' within the Crustacea, making the crustaceans paraphyletic<sup>3</sup> — that is, the group does not include all the descendants of its oldest ancestor ('dinosaurs' is a paraphyletic group because it does not include birds, which are derived from a dinosaur ancestor). The relative positions of other subphyla are still unresolved. To address this issue, Hwang *et al.*<sup>4</sup> sequenced the complete mitochondrial genome from the centipede *Lithobius forficulatus*, the first from a myriapod. Mitochondrial genomes are useful for tackling evolutionary events early in life's history ('deep phylogeny') because both the sequences and the relative order of protein-



**Figure 1 An embarrassment of arthropod morphology.** The extant arthropods come in various shapes and sizes. But they fall into five well-recognized groups or subphyla, whose evolutionary relationships are now being tackled by molecular and morphological analyses. Two conflicting hypotheses have been derived from their morphology and DNA sequence (Giribet *et al.*<sup>3</sup>, left) and mitochondrial genome sequence (Hwang *et al.*<sup>4</sup>, right). The principal differences are that Hwang *et al.* find evidence for a grouping of chelicerates and myriapods, whereas Giribet *et al.* find a comb-like pattern in which myriapods are more closely related to crustaceans and arthropods. Onychophorans, tardigrades and nematodes are related phyla used by Giribet *et al.* to root the arthropod tree; the analysis of Hwang *et al.* was rooted with more distant phyla. (Photographs, top to bottom, from: Spencer Entomological Museum, UBC; British Marine Life Study Society; Univ. Barcelona; Univ. California Berkeley, Museum of Paleontology; Erling Svensen; H. Ruhberg and H. Bosch; Mark Blaxter; M. Mundo-Ocampo *et al.*)

coding, ribosomal RNA and transfer RNA genes can be used.

Hwang and colleagues' mitochondrial data are consistent with a close link between the Myriapoda and Chelicerata, with the Hexapoda nested inside Crustacea. But the analyses are rooted with taxa that are distant relatives of the arthropods, because no complete onychophoran or tardigrade mitochondrial genomes have been published, and nematode mitochondria have a gene order and sequence that is too divergent to be used. The sequencing of pycnogonid, tardigrade and onychophoran mitochondrial genomes, currently underway, will be very useful for this sort of approach. The analysis of Hox gene sequences from a selection of arthropods and an onychophoran by Cook *et al.*<sup>3</sup> also supports a chelicerate–myriapod linkage, but these analyses lacked crustacean taxa (and may suffer from errors of inclusiveness).

However, two congruent data sets do not make a robust phylogeny. In the second of the papers in this issue, Giribet *et al.*<sup>5</sup> describe how they looked at the arthropods using sequence data from eight nuclear and two mitochondrial genes, and information on 303 morphological characters. By grouping data from individual species into higher groups, they were able to analyse 48 taxa, including pycnogonid sea spiders, as well as

onychophorans and tardigrades. Their finding is that the Chelicerata and Myriapoda are not sister taxa, and that 'mandibulate' arthropods (hexapods, crustaceans and myriapods) form a monophyletic group. The pycnogonids fell to the base of the arthropods, suggesting that sea-spider sequences and morphology may be useful for rooting the rest of the arthropod tree. The

only evidence for crustacean paraphyly was a grouping of *Drosophila* (fruitfly), *Daphnia* (water flea) and barnacles, but this appears to be an artefact resulting from a higher rate of change in the nuclear genes of these taxa.

It was never going to be easy to expose the true order of a series of events that happened more than 540 million years ago and left no obvious physical traces. From the initial analyses of single genes, the field has now progressed to using whole mitochondrial genomes and large combined data sets of morphology and nucleic-acid sequence from a full taxonomic range of species. Yet more data are still needed. There is only one true evolutionary tree, and future work will need to address the discrepancies between the reports of Hwang *et al.*<sup>4</sup> and Giribet *et al.*<sup>5</sup>. Nonetheless, the integrated approach used by these authors will not only drive the understanding of arthropod relationships, but will also help to resolve other long-standing problems in deep phylogeny, such as the relationships between the 'simple' animals (jellyfish and related organisms) and the 'higher' multicellular creatures. ■

Mark Blaxter is at the Institute of Cell, Animal and Population Biology, King's Buildings, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK.

e-mail: mark.blaxter@ed.ac.uk

1. Fortey, R. A. & Thomas, R. H. (eds) *Arthropod Relationships* (Chapman & Hall, London, 1998).
2. Peterson, K. J. & Eernisse, D. J. *Evol. Dev.* **3**, 170–205 (2001).
3. Cook, C. E., Smith, M. L., Telford, M. J., Bastianello, A. & Akam, M. *Curr. Biol.* **11**, 759–763 (2001).
4. Hwang, U. W., Freidrich, M., Tautz, D., Park, C. J. & Kim, W. *Nature* **413**, 154–157 (2001).
5. Giribet, G., Edgecombe, G. D. & Wheeler, W. C. *Nature* **413**, 157–161 (2001).
6. Akam, M. *Proc. Natl Acad. Sci. USA* **97**, 4438–4441 (2000).
7. Fortey, R. *Science* **293**, 438–439 (2001).
8. Bromham, L., Rambaut, A., Fortey, R., Cooper, A. & Penny, D. *Proc. Natl Acad. Sci. USA* **95**, 12386–12389 (1998).
9. Feng, D.-F., Cho, G. & Doolittle, R. F. *Proc. Natl Acad. Sci. USA* **94**, 13028–13033 (1997).
10. Jenner, R. A. *Evol. Dev.* **2**, 208–221 (2001).
11. Aguinaldo, A. M. *et al.* *Nature* **387**, 489–493 (1997).

## Nuclear physics

# Sizing up the heavyweights

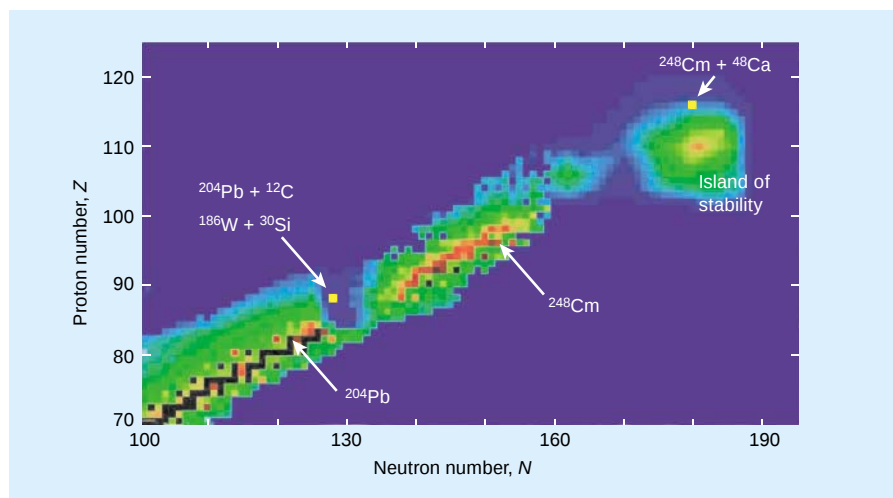
Yuri Oganessian

Colliding a heavy projectile with an even heavier target nucleus only occasionally produces superheavy elements. Analyses of the processes that prevent fusion suggest that projectile size is one of the problems.

A fundamental prediction of modern nuclear theory is the existence of an 'island of stability' among the largely unstable superheavy elements. By 1997 physicists had synthesized 12 new elements beyond fermium<sup>1</sup> (proton number,  $Z = 100$ ) that seemed to reinforce this intriguing hypothesis. During 1999 to 2001, new nuclides with large numbers of protons ( $Z = 114$  and  $116$ ) and neutrons ( $N = 174$  to  $176$ ) were synthesized<sup>2–4</sup> by colliding

neutron-rich ions with heavy target nuclei. This resulted in a considerable increase in the stability of nuclei with mass number  $A$  ( $A = N + Z$ ) of about 290, suggesting that the island of stability was not far away (Fig. 1). But superheavy elements are extremely difficult to synthesize because only one collision in  $10^{10}$  generates a new nuclide with the highest mass. On page 144 of this issue, Berriman *et al.*<sup>5</sup> offer theoretical and experimental analysis of the mechanisms underlying the





**Figure 1 The island of stability.** Red and black squares represent stable and long-lived nuclei; green squares represent artificially created nuclei, which decay shortly after formation. The probability of forming superheavy nuclei decreases as the island of stability, centred on  $Z = 114$  and  $N = 184$ , is approached. The heaviest known element,  $Z = 116$ ,  $N = 176$ , is formed by the reaction of  $^{248}\text{Cm}$  and  $^{48}\text{Ca}$ , and has a lifetime of 50 milliseconds. Its decay products,  $Z = 114$  and  $112$ , have lifetimes of about 3 seconds and 1 minute, respectively. The lower mass element studied by Berriman *et al.*<sup>5</sup>,  $^{216}\text{Ra}$ , can be generated by three different fusion reactions, including  $^{204}\text{Pb} + ^{12}\text{C}$  and  $^{186}\text{W} + ^{30}\text{Si}$ . Analysis and comparison of these three reactions provides a better understanding of the complex decay processes that prevent the fusion of heavy nuclei.

fusion and fission of massive nuclei. Their results provide insight into the complex processes that lead to the formation of super-heavy elements.

Artificial heavy nuclei typically decay less than a millisecond after formation. There are several decay mechanisms, but  $\alpha$ -decay (the emission of a  $^4\text{He}$  nucleus) and nuclear fission (by which the nucleus breaks into two fragments) are largely responsible. But this is just part of the problem. There are also difficulties associated with forming the compound nucleus in the first place. Niels Bohr first described nuclear fusion dynamics<sup>6</sup> in terms of an intermediate quasi-stable state (compound nucleus) to explain the fission of uranium atoms by bombarding them with slow neutrons. But Bohr did not take into account the time taken to create the compound nucleus, considering this stage to be 'instant', compared with the relatively long lifetime of the newly formed nucleus, which eventually decays through fission or by the emission of  $\gamma$ -rays.

This assumption is valid for neutrons but not for projectiles 50 times heavier. Moreover, if the projectile contains protons as well as neutrons, electrical (Coulomb) repulsion of bodies with the same charge may prevent the nuclei from fusing. Evidently, in the fusion of massive nuclei, the newly formed compound nucleus has a 'history' in which the nuclear system evolves until it reaches its final configuration. Such large-scale changes take a certain amount of time that appears to be comparable with the lifetime of the compound nucleus. Also, there is no guarantee that the complex collective motion associated with a nucleus consisting of hundreds of

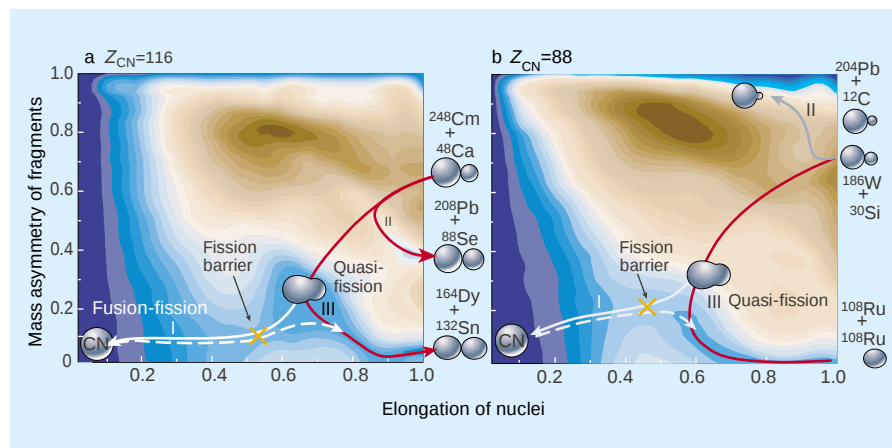
neutrons and protons will lead to the desired product rather than some other outcome. Such many-body problems are difficult to solve theoretically. But experimentalists also face complex problems. The decay signature of a rare superheavy nucleus has to be singled out from the enormous background of products created in unrelated reactions.

Consider the complete fusion of two nuclei in the formation of the compound

nucleus with  $Z = 116$  ( $N = 176$ ), the heaviest element yet created. Calculating the potential energy for all the possible configurations of this nucleus produces a potential-energy map for the formation of  $Z = 116$  (Fig. 2a). The evolution of this reaction is not straightforward: in the early stages, the nucleus can separate into two fragments with masses and charges similar to  $^{208}\text{Pb}$  and  $^{88}\text{Se}$ . This process is known as 'quasi-fission', as opposed to the genuine fission of the compound nucleus that produces two fragments of more comparable mass. Measurements<sup>7</sup> show that quasi-fission is much more probable in this reaction.

Nonetheless, the experimental observation of fission fragments with equivalent masses does not mean that they originate in the compound nucleus. As Fig. 2a shows, the composite system can also end up in the fission valley without first forming the compound nucleus (trajectory III). Separating out the contributions from these two processes is complicated because formation of the compound nucleus is much less probable than quasi-fission. So our understanding of the importance of quasi-fission and its role in preventing fusion is limited. An even thornier problem is getting information on the evaporation residues — the single nuclei that survive fission by the loss of neutrons from element 116.

Berriman *et al.*<sup>5</sup> take a different approach to the problem of quasi-fission. They study fusion reactions in which compound-nucleus formation is the most prominent process, and the competing reactions are much weaker. They consider a different fusion reaction



**Figure 2 Potential-energy surface maps for heavy nuclei created by fusion.** Higher energy values form hills (red-brown), whereas lower values form valleys (blue). The shape of the nuclei can change (in terms of the mass asymmetry of the fragments or the length of the nuclei), but a stable configuration, corresponding to a spherical compound nucleus (CN), must be reached. a, Map for a nucleus with  $Z = 116$  and  $N = 180$ , produced in the reaction  $^{248}\text{Cm} + ^{48}\text{Ca}$ . This compound nucleus can be reached along trajectory I, but then fissions into two fragments when it reaches the top of the fission barrier along the dash-marked trajectory (yellow cross). Trajectories II and III correspond to quasi-fission. b, Map for a nucleus with  $Z = 88$  and  $N = 128$ , produced in fusion reactions  $^{204}\text{Pb} + ^{12}\text{C}$  and  $^{186}\text{W} + ^{30}\text{Si}$ . For  $^{204}\text{Pb} + ^{12}\text{C}$ , the reaction quickly leads to the compound nucleus  $^{216}\text{Ra}$  without any possibility of quasi-fission. For  $^{186}\text{W} + ^{30}\text{Si}$ , the compound nucleus can be reached along trajectories I and II. In the first case there is a low probability of quasi-fission without forming the compound nucleus (trajectory III).

that results in a much lighter compound nucleus,  $^{216}\text{Ra}$  ( $Z = 88$ ,  $N = 128$ ), but has the advantage that three different reaction partners lead to the same nucleus. This means that the contribution of quasi-fission can be determined by measuring the probability of producing fission fragments and evaporation residues (isotopes of Ra with masses less than 216) simultaneously. But this raises another problem — separating the quasi-fission fragments from those of the compound-nucleus fission, which in this case are much more abundant.

To solve this problem, Berriman *et al.* compared three fusion reactions that lead to  $^{216}\text{Ra}$ , one of which doesn't involve quasi-fission at all. In this reaction the projectile  $^{12}\text{C}$  is swallowed up by the heavier nucleus,  $^{204}\text{Pb}$ , quickly resulting in the compound nucleus  $^{216}\text{Ra}$  (Fig. 2b). The authors identify the contribution of quasi-fission to the reaction dynamics by comparing the mass distributions of fission fragments for each reaction. In the absence of quasi-fission, fragment mass distributions from the same compound nucleus,  $^{216}\text{Ra}$ , should be identical. In the case of  $^{186}\text{W} + ^{30}\text{Si}$ , the small difference in the mass distributions is attributed to quasi-

fission, a process not expected to happen with a light projectile such as  $^{30}\text{Si}$ . Moreover, the authors show that the probability of fission increases with projectile mass, perhaps explaining the low success rates of experiments that use heavy projectiles to generate superheavy elements.

Here we have considered two extreme cases of formation and decay of nuclei with  $Z = 88$  and  $Z = 116$ , in which the probability of compound-nucleus formation differs by many orders of magnitude. Berriman *et al.*'s approach can be extended to a wider range of heavy nuclei than previously studied, and should stimulate further experimental and theoretical investigations of the processes behind superheavy fusion reactions. ■

Yuri Oganessian is in the Flerov Laboratory of Nuclear Reactions, Joint Institute for Nuclear Research, 141980 Dubna, Moscow, Russia.

e-mail: oganessian@flnr.jinr.ru

1. Hofmann, S. *Rep. Prog. Phys.* **61**, 639–689 (1988).
2. Oganessian, Yu. Ts. *et al. Nature* **400**, 242–245 (1999).
3. Oganessian, Yu. Ts. *et al. Phys. Rev. C* **62**, 041604(R) (2000).
4. Oganessian, Yu. Ts. *et al. Phys. Rev. C* **63**, 011301(R) (2000).
5. Berriman, A. C. *et al. Nature* **413**, 144–147 (2001).
6. Bohr, N. *Nature* **137**, 344–348 (1936).
7. Itkis, M. G. *et al. in Proc. Second Int. Conf. Fission and Properties of the Neutron-Rich Nuclei* 268–285 (World Scientific, Singapore, 1999).

## Metabolism

# Controlling the glucose factory

Antonio Vidal-Puig and Stephen O'Rahilly

In times of starvation the liver turns into a glucose-producing organ, providing fuel for the brain. The hormonal signals that control this switch in glucose metabolism may converge on a single regulatory molecule.

Mammals have evolved sophisticated systems to maintain glucose levels in the blood within tight limits, despite large fluctuations in food intake. The main drive for such strict control is the brain's near-absolute requirement for glucose as an energy source. So the ability of the liver and, to a lesser extent, the kidneys to convert stored fuels into glucose is essential to surviving even short periods of starvation. The production of glucose by the liver is finely tuned by hormonal signals, which influence both the breakdown of glycogen stores to produce glucose and the synthesis of glucose from carbon-based precursors<sup>1</sup>. But how are these signals transmitted and coordinated within liver cells? On pages 131 and 179 of this issue, Yoon *et al.*<sup>2</sup> and Herzig *et al.*<sup>3</sup> provide compelling evidence for the involvement of a protein known as PGC-1. This is surprising because the liver was not thought to be a major site of expression of this molecule.

PGC-1 was first identified<sup>4</sup> three years ago as a new coactivator-like molecule that interacts with PPAR- $\gamma$ , a hormone receptor found in the nucleus of fat cells. Coactivators

are key parts of the multiprotein complexes that control the transcription of genes into messenger RNA. They do not bind DNA themselves but instead provide a functional link between the proteins (such as nuclear hormone receptors) that bind the regulatory elements of genes, and the process of mRNA synthesis. PGC-1 is known<sup>4,5</sup> to have a role in coordinating adaptive thermogenesis — the generation of heat in response to cold or eating high-fat foods — at the transcriptional level. But it may have many more functions. For example, it works with a muscle-specific transcription factor to regulate glucose transport<sup>6</sup>, and it might mediate some of the transcriptional effects of the oestrogen receptor<sup>7</sup>.

It is not yet known exactly how PGC-1 works. It has been reported to promote transcription by assembling a complex containing histone acetyltransferases, which control the access of transcription factors to DNA, and its ability to activate the acetyltransferases might depend on its undergoing a conformational change, induced by binding to nuclear receptors such as PPAR- $\gamma$ <sup>8</sup>. PGC-1 can also alter RNA processing

when it is part of a transcriptionally active complex<sup>9</sup>.

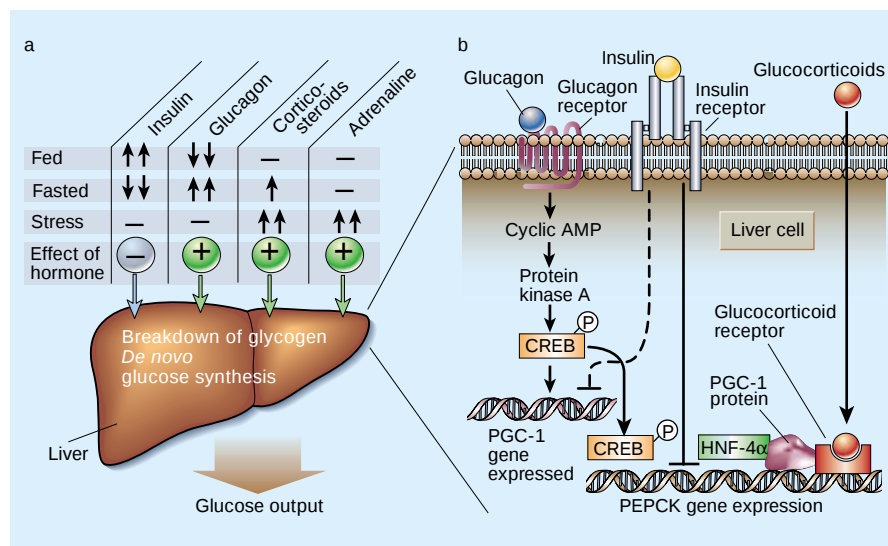
Yoon *et al.*<sup>2</sup> and Herzig *et al.*<sup>3</sup> do not address the question of how PGC-1 works, but rather look again at what it does. Specifically, they have investigated whether this protein is involved in the control of blood glucose levels by the liver. When mammals are starved for a short time, circulating levels of insulin drop and levels of glucagon and corticosteroids rise (Fig. 1a, overleaf). One eventual result is the generation of glucose from carbon precursors in the liver, a process known as gluconeogenesis. The key initial observation underlying the new studies was Yoon *et al.*'s discovery<sup>2</sup> that although, as previously reported, PGC-1 mRNA is not significantly expressed in livers from fed rats, it is readily detected in livers from food-deprived animals. The authors guessed that PGC-1 might have an active role in controlling gluconeogenesis in the liver in response to starvation.

Supporting this idea, Yoon *et al.*<sup>2</sup> find that levels of PGC-1 mRNA are high in the liver in three different rodent models of increased liver gluconeogenesis. Moreover, overexpression of PGC-1 in liver cells — in culture and in whole animals — increases glucose production and the transcription of genes encoding gluconeogenic enzymes. At one of these genes, PGC-1 works with two key transcription factors: hepatic nuclear factor-4 $\alpha$  and the glucocorticoid receptor (Fig. 1b).

How is the expression of PGC-1 upregulated in response to starvation? It has long been known that hormones such as glucagon and adrenaline, which increase the levels of the messenger molecule cyclic AMP in the liver, tend to counteract the effects of insulin and promote glucose release from the liver. Herzig *et al.*<sup>3</sup> provide evidence that the metabolic effects of cAMP in the liver may be mediated largely through PGC-1. The authors looked at three different mouse models in which cAMP signalling in the liver is impaired. In all three models, levels of blood glucose were very low and there was a marked reduction in the expression of gluconeogenic enzymes and PGC-1 in the liver. Yoon *et al.*, meanwhile, report that levels of PGC-1 mRNA can be increased in liver cells by exposure to cAMP analogues. Herzig *et al.* show that the regulatory region of the PGC-1 gene contains a functional cAMP-response element, and that transcription from this region is inhibited by insulin.

These results indicate that PGC-1 is a key protein involved in coordinating the liver's response to fasting, although formal proof that it is essential for the normal control of gluconeogenesis will have to await studies of animals in which the PGC-1 gene has been deleted from the liver. Meanwhile, the new studies<sup>2,3</sup> may have broader implications for understanding glucose metabolism. Insulin is usually considered to be the dominant





**Figure 1** Glucose production in the liver. **a**, Hormones that regulate the liver's glucose output, and their fluctuation (short arrows) in different circumstances. **b**, Molecular details of the processes shown in **a**. Yoon *et al.*<sup>2</sup> and Herzig *et al.*<sup>3</sup> find that the protein PGC-1 is a key regulator of gluconeogenesis (glucose synthesis). It works with transcription factors, including hepatic nuclear factor-4 $\alpha$  (HNF-4 $\alpha$ ) and glucocorticoid receptors, to activate the expression of gluconeogenic genes, such as that encoding phosphoenolpyruvate carboxykinase (PEPCK). Left, glucagon activates the cyclic-AMP signalling pathway, which increases glucose production through a mechanism that involves the expression of PGC-1. Centre, insulin counteracts gluconeogenesis by blocking the expression of genes such as the PEPCK gene, perhaps by inhibiting PGC-1 expression (dashed line). Right, glucocorticoids enter cells and bind to receptors, which, with PGC-1, activate the expression of gluconeogenic genes.

regulator of glucose output from the liver, with hormones like glucagon, adrenaline and glucocorticoids having a secondary role, particularly under stress. For example, after the pancreas is removed — depriving the liver of both insulin and glucagon — the liver responds by massively increasing glucose production because of the lack of insulin, not by shutting off glucose production because of a glucagon shortage<sup>10</sup>.

So it is intriguing that interfering with liver cAMP-mediated pathways alone can lead to such low levels of blood glucose in intact organisms<sup>3</sup>: one might expect that a reduction in insulin secretion, occurring in response to small decreases in blood glucose, would compensate. In the mouse models studied by Herzig *et al.*, insulin levels do indeed drop a little, but not enough to compensate for the loss of cAMP signalling. This should make us think more deeply about which signals, other than glucagon and adrenaline, might be important in maintaining 'physiological' cAMP signalling in the liver. Another unanswered question is the extent to which alterations in PGC-1 expression can explain the suppressive effects of insulin on glucose production.

The new observations may also have relevance to human disease. Diabetes mellitus, particularly when poorly controlled, is characterized by increased gluconeogenesis in the liver<sup>11</sup>. One of the main treatments for the most common form of diabetes is metformin, which works mainly by reducing

glucose production by the liver<sup>12</sup>. But this drug is not without side effects, and its mechanism of action is unknown. A better knowledge of the molecules that control gluconeogenesis might allow the design of more focused drugs.

Finally, perhaps the most important aspect of these papers<sup>2,3</sup> is their illustration of how a range of hormonal signals can regulate a metabolic process by altering the levels of a coactivator protein. Coactivators interact indirectly with many genes through many transcription factors, making this an economical way of orchestrating a complex transcriptional response. It seems unlikely that PGC-1 is an isolated example of this strategy.

Antonio Vidal-Puig and Stephen O'Rahilly are in the Departments of Medicine and Clinical Biochemistry, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK. e-mail: sorahill@hgmpr.mrc.ac.uk

1. Tappy, L. & Minehira, K. *Curr. Opin. Clin. Nutr. Metab. Care* **4**, 273–277 (2001).
2. Yoon, J. C. *et al.* *Nature* **413**, 131–138 (2001).
3. Herzig, S. *et al.* *Nature* **413**, 179–183 (2001).
4. Puigserver, P. *et al.* *Cell* **92**, 829–839 (1998).
5. Wu, Z. *et al.* *Cell* **98**, 115–124 (1999).
6. Michael, L. F. *et al.* *Proc. Natl Acad. Sci. USA* **98**, 3820–3825 (2001).
7. Tcherepanova, I., Puigserver, P., Norris, J. D., Spiegelman, B. M. & McDonnell, D. P. *J. Biol. Chem.* **275**, 16302–16308 (2000).
8. Puigserver, P. *et al.* *Science* **286**, 1368–1371 (1999).
9. Monsalve, M. *et al.* *Mol. Cell* **6**, 307–316 (2000).
10. Giacca, A., Fisher, S. J., McCall, R. H., Shi, Z. Q. & Vranic, M. *Endocrinology* **138**, 999–1007 (1997).
11. Mitrakou, A. *et al.* *New Engl. J. Med.* **326**, 22–29 (1992).
12. DeFronzo, R. A. *Ann. Intern. Med.* **131**, 281–303 (1999).

## Daedalus

### Extremes of rubbish

**Burial or incineration?** This dilemma applies not only to your mortal remains, but to all the rubbish you discard throughout life. Incineration promises (but seldom delivers) an immediate output of useful thermal power. Landfill, curiously, also offers energy, but more slowly. Most rubbish dumps slowly evolve methane, from bacterial fermentation of their biochemical content (food residues, paper, disposable nappies and so on). The gas is a major nuisance — it makes old landfill sites hazardous to build on, for example — but there are schemes to exploit it as a fuel.

As a fermenter, says Daedalus, a landfill dump is absurdly big, runs absurdly slowly, and degrades only a tiny fraction of its contents. The organisms that inhabit it are simply not up to the job. To be economic, it should be much smaller, with a much faster throughput. Daedalus recalls the 'extremophile' bacteria found in hot springs, volcanic black smokers and similar thermal and chemical extremes. They perform the most amazing chemistry, and at speeds accelerated by the high temperature. So DREADCO biochemists are now playing with promising extremophiles — swapping their genes around, exposing them to radiation, selecting the survivors of really brutal conditions, and so on — seeking the ideal cultural mix for DREADCO's Extreme Rubbish Fermenter, or ERF. It will challenge the accepted limits of the conditions under which life can survive.

Daedalus cannot guess what those limits will be. He hopes that ERF will operate at 100 °C at least; with luck it will even go superheated. The harsher the conditions, the more types of rubbish will become fermentable. Not only will plastics and other synthetics soften, dissolve and decompose, becoming food for the voracious extremophiles, but even metals may begin to succumb. Only glass and ceramics (of which ERF itself will be made) will remain aloof.

ERF will resemble an intense, high-pressure, giant variant of the garden compost-bin. Rubbish will be rammed into the top, fuel gas will be extracted from peripheral pipes, and the sterile residue, a sort of rusty grit, will emerge at the bottom. It will be self-heating. Its ferocious internal conditions will impose intense evolutionary pressure on the extremophiles inside. They should steadily evolve further, perhaps imitating the way that life itself developed in the hot and chemically aggressive environment of the young Earth.

David Jones

# Pack formation in cycling and orienteering

Avoiding conditions that draw competitors together may be a better way to test ability.

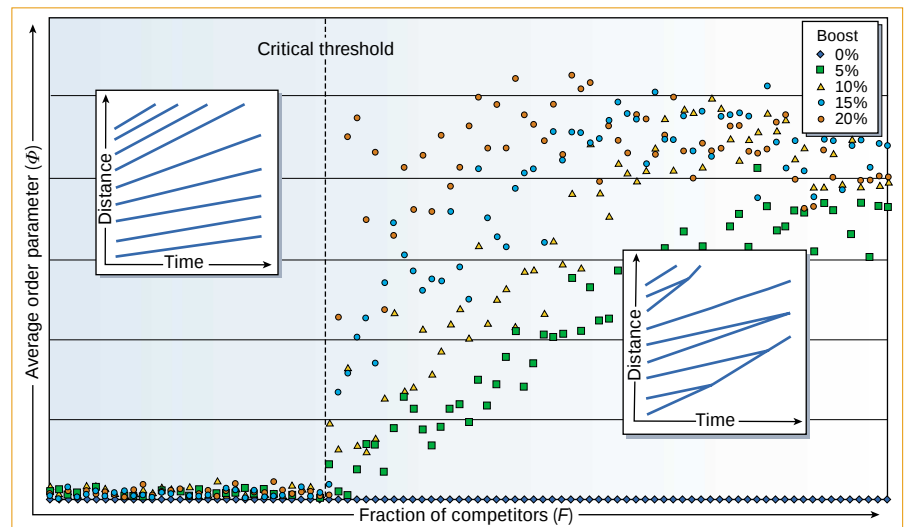
In cycling and orienteering competitions, competitors can become bunched into packs, which may mask an individual's true ability. Here we model this process with a view to determining when competitors' times are determined more by others than by their own ability. Our results may prove useful in helping to stage events so that pack formation can be avoided.

We integrated equations of motion in discrete time steps for interacting competitors moving in one dimension from start to finish, with passing allowed. Individually, the  $i$ th competitor moves at speed  $u_i$ , which increases by  $u_i \times b$  (where  $b$  represents the 'boost' factor) when another competitor is within a certain range ahead.

The appropriate percentage value for the boost and length of the range will be different for different sports, but the same for different competition formats in the same sport. In cycling (Fig. 1), the boost is due to aerodynamic and psychological effects<sup>1</sup>, whereas in orienteering it occurs when a competitor is able to follow another within sight without the need to map-read. In each case, the leading competitor is unaffected. The number of competitors, distribution of  $u_i$  values, and initial separations depend on the particular competition



**Figure 1** Go with the flow: in cycling races, competitors often bunch into packs, giving them an aerodynamic advantage which may belie an individual's true ability.



**Figure 2** For quantitative study, we define an order parameter,  $\Phi$ , which increases with packing: for  $N+1$  competitors,  $\Phi(t) = -N^{-1} \log \Pi_{j,j+1}(t)/2r_0 + 1/2$ , where  $r_{j,j+1}(t)$  is the separation between the  $j$ th and  $j+1$ th competitors and  $r_0$  is the mean separation. Competitors with  $u_i$  values that vary by more than the boost ( $b$ ) cannot stay together, so the maximum value of  $\Phi$  depends on the speed distribution. The main graph shows values of  $\Phi$  with increasing  $F$  (as determined by setting  $b=0$ ), from 500 chasing-start races, each with 30 competitors. In these races, competitors are started in order of highest  $u_i = u_i(1+R)$ , where  $R$  is a random number between 0 and 0.1. Insets, details from simulations showing typical competitor behaviour in diverging (left, low  $F$ ) and pack-forming (right, high  $F$ ) regimes. Qualitative results are stable against noise (imposed as time-dependent fluctuations in  $u_i$ ) up to the level of  $b$ .

and may vary between races. For mass starts (zero initial separation), maximal bunching is seen to occur immediately. In time trials (constant initial separation, with  $u_i$  distributed randomly), packs are ultimately formed, but often only after a considerable time.

Interesting cases are chasing starts, in which competitors are approximately sorted by  $u_i$ . For example, competitors' cycling start times in triathlons are determined by swim time (as swimming immediately precedes cycling), which is correlated with cycling ability, and start times in some orienteering events are determined by a preliminary race. Here packs are formed after a short time, if at all. After a longer time, these packs (or individuals) move steadily apart. This reveals a sharp crossover between packing and non-packing behaviour: below a critical ratio of range to starting interval, most competitors move individually and move steadily further apart. Above this critical ratio, however, packs form that catch and absorb individuals.

Surprisingly, the position of the crossover depends on a single criterion: the fraction ( $F$ ) of competitors who have abilities within the boost factor and who are able to get within range of others in the absence of boost interactions. If  $F > 13\%$  (the critical threshold; Fig. 2), all competitors are rapidly swept into packs. In time

trials, the onset of pack formation coincides with the time required to meet this criterion;  $b$  and  $u_i$  determine pack size.

Pack formation is stable against random fluctuations of  $u_i$  at each time point, provided that these are less than  $b$ , and is unchanged by adding a further boost that is proportional to pack size, or by making  $b$  range-dependent.

Data are available from orienteering races<sup>2</sup> in which competitors are electronically timed at a series of checkpoints. Our position/time simulation data compare well qualitatively to actual races, and the  $F = 13\%$  threshold distinguishes those races in which pack formation dominates. Our results seem to be qualitatively similar to the formation of shockwaves or traffic jams<sup>3</sup>, but an important difference is that our packs tend to be composed of the same competitors throughout.

We have derived the conditions for pack formation: packs form in chasing and time-trial races when more than 13% of competitors encounter others who are then able to stay with them because of the benefits of following.

**G. J. Ackland, D. Butler**

Department of Physics and Astronomy,  
University of Edinburgh, Edinburgh EH9 3JZ, UK  
e-mail: g.j.ackland@ed.ac.uk

1. Olds, T. *Eur. J. Appl. Physiol.* **77**, 492–497 (1998).
2. British Orienteering Federation results.  
<http://www.cix.co.uk/~bof> (2001).
3. Nagel, K. & Schreckenberg, M. *J. Physique* **2**, 2221–2229 (1992).



Sexual selection

# Are ducks impressed by drakes' display?

Surprisingly few birds have penises, but among those that do, the Argentine lake duck (*Oxyura vittata*) tops the bill — the penis of this small stiff-tail duck from South America is shaped like a corkscrew and, at almost half a metre long, is the largest of any bird measured so far. Factors responsible for the evolution of this remarkable organ could include runaway selection, whereby drakes with longer penises gain dominance and copulate with more females, or preference by females for drakes with longer and more decorated penises.

This small duck (which typically weighs around 640 g) was previously reported to have a 20-cm penis<sup>1,2</sup>, which is about half of the drake's total body length and is comparable to the penis of an ostrich (*Struthio camelus*)<sup>3</sup>. A new voucher specimen from Argentina's Córdoba province (Fig. 1) indicates, however, that the penis size of this species was underestimated. We measured the length of the unwound, everted penis of an aroused *O. vittata* drake and found that it was 32.5 cm long when hanging under the force of gravity, but that it could stretch

when fully unwound to 42.5 cm.

Like other stiff-tail ducks, lake ducks are promiscuous and boisterous in their sexual activity<sup>4</sup>. One explanation for the evolution of such a long penis might be selection for its competitive advantage — for instance, groups of drakes might display their everted penises to attract females. Success in sperm competition<sup>5</sup> may also be a factor. The base of the lake-duck penis is covered with spines, yet the tip is soft and brush-like. Before ejaculation, drakes probably use their penises like bottle-brushes to remove sperm stored in the oviduct by the female's previous consort. The larger the bottle-brush, the more effective it will be, clearing the way for the drake to father the ducklings.

Many questions remain unanswered. How much of his penis does the drake actually insert, and does the anatomy of the female's oviducts make them unusually difficult to inseminate? The Argentine lake duck offers a sizeable opportunity to study sexual selection and sperm competition in birds.

**Kevin G. McCracken\*†, Robert E. Wilson\*, Pamela J. McCracken†, Kevin P. Johnson‡**

\*Institute of Arctic Biology and Department of Biology and Wildlife, University of Alaska Fairbanks, Fairbanks, Alaska 99775, USA  
e-mail: fnkgm@uaf.edu

†University of Alaska Museum, Fairbanks, Alaska 99775, USA

‡Illinois Natural History Survey, Champaign, Illinois 61820, USA

1. McCracken, K. G. *Auk* **117**, 820–825 (2000).
2. Mead, C. *Nature* **407**, 461 (2000).
3. Briskie, J. V. & Montgomerie, R. J. *Avian Biol.* **28**, 73–86 (1997).
4. Johnsgard, P. A. & Carbonell, M. *Ruddy Ducks and Other Stiff-tails: Their Behavior and Biology* (Univ. Oklahoma Press, Norman, 1996).
5. Birkhead, T. R. & Møller, A. P. *Sperm Competition in Birds: Evolutionary Causes and Consequences* (Academic, London, 1992).

Nitrogen fixation

# Endocrine disrupters and flavonoid signalling

Nitrogen fixation is a symbiotic process initiated by chemical signals from legumes that are recognized by soil bacteria. Here we show that some endocrine-disrupting chemicals (EDCs)<sup>1–3</sup>, so called because of their effect on hormone-signalling pathways in animal cells, also interfere with the symbiotic signalling that leads to nitrogen fixation. Our results raise the possibility that these phytochemically activated pathways may have features in common with hormonal signalling in vertebrates, thereby extending the biological and ecological impact of EDCs.

Initiation and maintenance of nitrogen-fixing symbiosis between rhizobial bacteria (*Sinorhizobium meliloti*) and alfalfa (*Medicago sativa*) requires an exchange of signals

between these organisms<sup>4</sup>. In *S. meliloti*, the constitutively expressed product of the *nodD* genes acts as a positive regulator which, upon recognizing agonistic phytochemicals produced by alfalfa, activates transcription of the common nodulation (*nod*) genes needed to initiate symbiosis<sup>4,5</sup>. Whereas some flavonoids, such as luteolin and apigenin, stimulate *nodD* activation, others (such as chrysin) can antagonize the induction of *nod* genes<sup>5–7</sup>. Chemicals in the environment that mimic or interfere with this phytochemical signalling to *S. meliloti* may confound recognition crucial to symbiosis.

As luteolin and other flavonoids have been proposed to mimic oestrogen signalling in mammalian cells<sup>2</sup> and a possible homology has been identified between NodD and the oestrogen receptor ER- $\alpha$  (ref. 8), we investigated whether EDCs that disrupt oestrogen signalling could also interfere with symbiotic signalling<sup>3</sup>. We tested the effects in *S. meliloti* of different compounds, added in the presence of luteolin, on activation and transcription of a *nod* reporter gene using a  $\beta$ -galactosidase assay<sup>9</sup>.

The inhibitory flavonoids chrysin and genistein<sup>7,10</sup> were among the most potent phytochemical–NodD signalling inhibitors, as were the organochlorine pesticides pentachlorophenol and methyl parathion, which each reduced *nod* expression by up to 90% even in the presence of luteolin (Fig. 1a) or apigenin (results not shown).

Half-maximal inhibitory concentrations (IC<sub>50</sub>) were comparable for the inhibitory phytochemicals and the EDCs tested (Fig. 1a). Even nanomolar concentrations of some pesticides significantly inhibited phytochemical–NodD signalling (Fig. 1a). The insecticide dichlorodiphenyltrichloroethane (DDT) and its metabolites inhibited *nod* expression by an average of 40% (see, for example, Fig. 1b); other EDCs<sup>3</sup>, including the plasticizer bisphenol A (Fig. 1a) and the herbicides 2,4-D and 2,4,5-T (results not shown), reduced expression by 66%, 32% and 37%, respectively. Endosulfan, methoxychlor, aldrin, dieldrin, dursban, vinclozolin and diazinon had no effect on symbiotic signalling (results not shown).

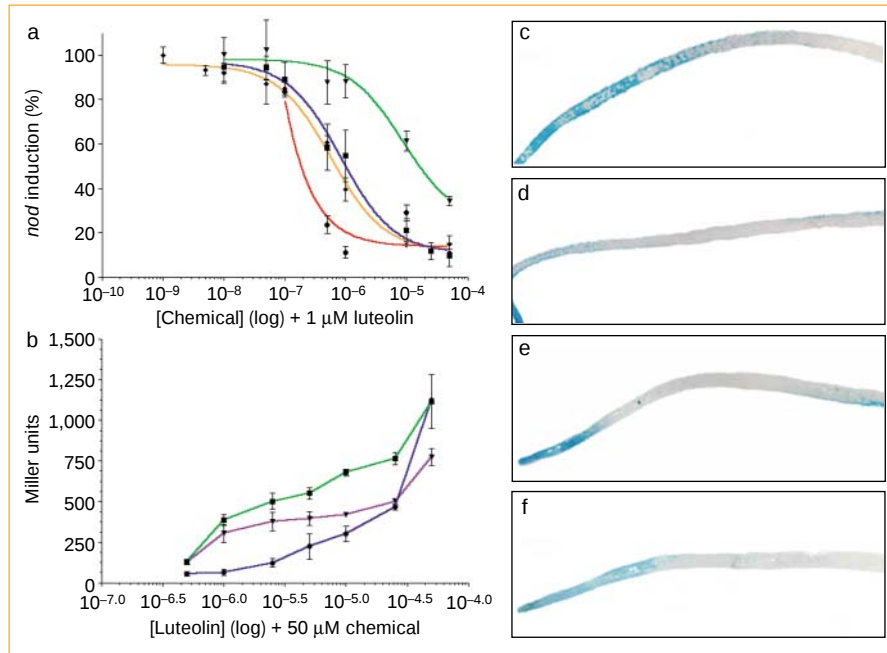
The inhibitory EDCs are all planar phenolic compounds, whereas those that had no effect in our assay are non-planar and have no free hydroxyl group. Inhibition of transcription regulated by the luteolin–NodD pathway can be overcome by increasing the concentration of luteolin (Fig. 1b).

We grew alfalfa seedlings inoculated with the bacterial reporter strain on medium containing X-gal<sup>11,12</sup>; NodD-initiated gene transcription produces  $\beta$ -galactosidase which, in the presence of X-gal, generates a blue product<sup>11</sup>. Figure 1c–f shows that  $\beta$ -galactosidase is expressed along the entire root system in the alfalfa control, whereas blue product is almost absent in roots treated with chrysin,



**Figure 1** Photograph of the 42.5-cm penis of a male Argentine lake duck (*Oxyura vittata*) taken near Arroyo Chucul, Córdoba, Argentina on 30 April 2001.





**Figure 1** Endocrine-disrupting chemicals (EDCs) alter symbiotic signalling *in vitro* and *in vivo*. **a**, Phytochemicals (chrysin, orange) or EDCs (methyl parathion, green; pentachlorophenol, blue; bisphenol A, red) at 50 nM to 50 μM were added to *Sinorhizobium meliloti* 1021pRmM57 (ref. 9; provided by S. R. Long) treated with 1 μM luteolin. Luteolin alone was set as 100% induction<sup>10</sup>; luteolin plus chemical was compared to 100% induction. IC<sub>50</sub> values (molar): methyl parathion,  $4.3 \times 10^{-7}$ ; pentachlorophenol,  $9.9 \times 10^{-7}$ ; chrysin,  $7.0 \times 10^{-7}$ ; bisphenol A,  $1.7 \times 10^{-5}$ . **b**, 50 μM o,p'-DDT (purple) or pentachlorophenol (blue) caused 43% and 90% inhibition, respectively, of *nod* activation (measured as Miller units of β-galactosidase activity). Increasing the luteolin concentration (100 nM to 50 μM) restores *nod* activation to normal. **c–f**, Alfalfa seedlings were grown on buffered nodulation medium plus X-gal<sup>11,12</sup> and inoculated with dilute<sup>12</sup> *S. meliloti* 1021pRmM57 treated with vehicle (dimethylsulphoxide, **c**), chrysin (**d**), methyl parathion (**e**) or pentachlorophenol (**f**) at 50 μM. *NodC–lacZ* expression is visualized as blue product. Further details are available from the authors.

pentachlorophenol or methyl parathion, except between the tip and the first emerging root hairs, where luteolin concentration is highest<sup>6,10</sup>. This effect is similar to that shown in Fig. 1b, where inhibition is overcome by increasing concentrations of luteolin.

Nitrogen fixation is controlled by both agonistic and antagonistic phytochemical signalling to *Rhizobium*<sup>5,13</sup>. As with the vertebrate endocrine system, plant–bacterial symbiosis relies on recognition of specific agonists and antagonists for gene regulation<sup>2,10,13</sup>. We have shown that some pesticides and planar phenolic EDCs can interfere with plant–*Rhizobium* signalling and nitrogen-fixing symbiosis, thereby exposing a previously unrecognized similarity to the effects of EDCs on vertebrate endocrine signalling.

**Jennifer E. Fox\***†, **Marta Starcevic\***,  
**Kelvin Y. Kow\***, **Matthew E. Burow\***‡,  
**John A. McLachlan\***‡

\*Environmental Endocrinology Laboratory, Center for Bioenvironmental Research at Tulane and Xavier Universities, †Molecular and Cellular Biology Program, Tulane University, and ‡Department of Pharmacology, Tulane University Medical School, New Orleans, Louisiana 70112, USA  
e-mail: john.mclachlan@tulane.edu

1. Tham, D. M., Gardner, C. D. & Haskell, W. L. *J. Clin. Endocrinol. Metab.* **83**, 2223–2235 (1998).
2. Collins-Burrow, B. M. *et al. Nutr. Cancer* **38**, 229–244 (2000).
3. McLachlan, J. A. *Endocrinol. Rev.* **22**, 319–341 (2001).
4. Schultze, M. & Kondorosi, A. *Annu. Rev. Genet.* **32**, 33–57 (1998).
5. Peters, N. K., Frost, J. W. & Long, S. R. *Science* **233**,

- 977–980 (1986).
6. Djordjevic, M. A. *et al. EMBO J.* **6**, 1173–1179 (1987).
7. Firmin, J. L. *et al. Nature* **324**, 90–93 (1986).
8. Gyorgypal, Z. & Kondorosi, A. *Mol. Gen. Genet.* **226**, 337–340 (1991).
9. Mulligan, J. T. & Long, S. R. *Proc. Natl Acad. Sci. USA* **82**, 6609–6613 (1985).
10. Peters, N. K. & Long, S. R. *Plant Physiol.* **88**, 396–400 (1988).
11. Redmond, J. W. *et al. Nature* **323**, 632–635 (1986).
12. Gage, D. J., Bobo, T. & Long, S. R. *J. Bacteriol.* **178**, 7159–7166 (1996).
13. Baker, M. E. *Adv. Exp. Med. Biol.* **439**, 249–276 (1998).

#### COMMUNICATIONS ARISING

##### Palaeovegetation

## Diversity of temperate plants in east Asia

The exceptionally broad species diversity of vascular plant genera in east Asian temperate forests, compared with their sister taxa in North America, has been attributed to the greater climatic diversity of east Asia, combined with opportunities for allopatric speciation afforded by repeated fragmentation and coalescence of populations through Late Cenozoic ice-age cycles<sup>1</sup>. According to Qian and Ricklefs<sup>1</sup>, these opportunities occurred in east Asia because temperate forests extended across the continental shelf to link populations in China, Korea and

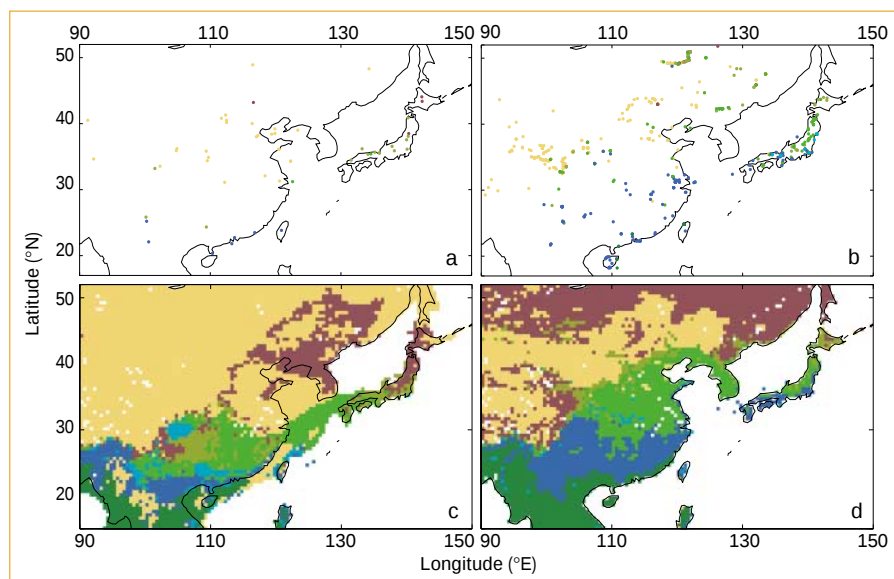
Japan during glacial periods, whereas higher sea levels during interglacial periods isolated these regions and warmer temperatures restricted temperate taxa to disjunct refuges. However, palaeovegetation data from east Asia<sup>2–6</sup> show that temperate forests were considerably less extensive than today during the Last Glacial Maximum, calling into question the coalescence of tree populations required by the hypothesis of Qian and Ricklefs<sup>1</sup>.

Pollen data<sup>2–6</sup> from <sup>14</sup>C-dated sediment cores, processed using a standard interpretive technique<sup>7</sup>, record the changes in vegetation between the Last Glacial Maximum (Fig. 1a) and the present (Fig. 1b). These data are derived from work carried out by Chinese and Japanese palynologists during the two decades since the reconstruction by CLIMAP cited in ref. 1. The data indicate that temperate forests were restricted in distribution during the Last Glacial Maximum, a finding that is contrary to the inference by Qian and Ricklefs<sup>1</sup>. The forests were displaced either by boreal and mixed forests lacking most of the temperate forest taxa, or by treeless vegetation (steppe and desert, with pollen assemblages dominated by *Artemisia*, *Chenopodiaceae* and grasses), over much of their present range.

Low levels of precipitation, compounded by the low water-use efficiency of C<sub>3</sub> photosynthesis at atmospheric CO<sub>2</sub> concentrations during the Last Glacial Maximum (less than 200 p.p.m.; refs 7, 8), confined temperate forests to medium elevations in northern China, whereas low temperatures caused mixed forests (mixtures of boreal conifers with only the most cold-tolerant temperate taxa) to dominate in much of Japan and southern China<sup>2–6,8</sup>.

To provide a spatially continuous vegetation map (Fig. 1c) that is physically consistent with climate forcing for the Last Glacial Maximum, we used a model of climate in the last glacial maximum (based on known changes in insolation, atmospheric composition and physiography)<sup>9</sup> as input to an equilibrium biogeography model, BIOME4 (Fig. 1d)<sup>10</sup>. The large-scale features of this simulation are supported by pollen data (Fig. 1a). Temperate forests are shown as confined to a relatively narrow belt at low elevations, whereas non-forest vegetation is shown to be greatly extended, as the pollen data indicate. On the other hand, a band of temperate deciduous forest is shown extending across the East China Sea, consistent both with Qian and Ricklefs' hypothesis<sup>1</sup> and with pollen evidence for the existence of an offshore glacial refuge for temperate deciduous trees.

Quaternary palaeodata have repeatedly overturned historical biogeographical hypotheses that were based on present-day taxon distributions. The palaeodata indicate a more complex history in which



**Figure 1** Observed and simulated vegetation of east Asia during the Last Glacial Maximum and today. **a**, Pollen data for the Last Glacial Maximum, 18,000 ± 1,000 <sup>14</sup>C yr BP, classified to top-level vegetation units (biomes) using a standard method<sup>1</sup>. **b**, Subrecent pollen data analysed in the same way, reflecting present vegetation patterns. **c**, Simulated vegetation during the Last Glacial Maximum, 21,000 cal yr BP (~18,000 <sup>14</sup>C yr BP). Climate anomalies from the NCAR CCM1, with computed (mixed-layer ocean model) sea-surface temperatures<sup>2</sup>; vegetation simulation from BIOME4 (ref. 10). **d**, Simulated present potential natural vegetation, on the basis of twentieth-century climatology (CLIMATE 2.2). Colour scale: dark green, tropical forest, savanna and woodland; light green, temperate deciduous forest; light blue, temperate coniferous forest; dark blue, warm temperate evergreen forest; olive green, mixed (temperate and boreal) forest; brown, boreal forest; orange, non-forest.

habitat reductions in some regions occurred simultaneously with increases in others. Palaeodata from east Asia may still be broadly consistent with the allopatric-speciation hypothesis for species diversity put forward by Qian and Ricklefs<sup>1</sup>, but only if it is recognized that population fragmentation occurred during glacial as well as interglacial periods. Although eastern North America also shows evidence of relative aridity during the Last Glacial Maximum, temperate forests there seem to have been more extensive and continuous<sup>7</sup>.

**S. P. Harrison\***†, **G. Yu\***‡, **H. Takahara§**, **I. C. Prentice\***

\*Max Planck Institute for Biogeochemistry, PO Box 100164, 07701 Jena, Germany  
e-mail: sharris@bgc-jena.mpg.de

†Dynamic Palaeoclimatology, Lund University, 22300 Lund, Sweden

‡Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210093, China  
§University Forest, Kyoto Prefectural University, Kyoto 606, Japan

1. Qian, H. & Ricklefs, R. E. *Nature* **407**, 180–182 (2000).
2. Yu, G. et al. *J. Biogeogr.* **27**, 635–664 (2000).
3. Takahara, H. et al. *J. Biogeogr.* **27**, 665–683 (2000).
4. Xu, J. S. *Selected Papers from the First Symposium of the Palynological Society of China* (Scenic, Beijing, 1982) (in Chinese).
5. Han, Y. S. & Meng, G. L. *Oceanogr. Limnol. China* **17**, 196–205 (1986) (in Chinese).
6. Meng, G. L. & Wang, S. Q. *Oceanogr. Limnol. China* **18**, 253–263 (1987) (in Chinese).
7. Prentice, I. C., Jolly, D. & BIOME 6000 Participants. *J. Biogeogr.* **27**, 507–519 (2000).
8. Farrera, I. et al. *Clim. Dyn.* **15**, 823–856 (1999).
9. Kutzbach, J. et al. *Q. Sci. Rev.* **17**, 473–506 (1998).
10. Kaplan, J. O. *Geophysical Applications of Vegetation Modeling*. Thesis, Univ. Lund (2001).

**Qian and Ricklefs reply** — The point we wished to make was simply that the more complex geography and topography of eastern Asia compared with eastern North America, in conjunction with climate change and sea-level fluctuations, have provided greater opportunity for allopatric speciation. This explanation of the greater diversity of vascular plants in temperate regions of eastern Asia cannot yet be tested using any particular biome reconstruction, all of which are poorly resolved.

The estimated distribution of vegetation during the Last Glacial Maximum (18,000 yr BP) and at 6,000 yr BP, on the basis of fossil-pollen data<sup>1</sup> and climate, shows that mesic vegetation types were shifted southwards during the Last Glacial Maximum relative to their modern distributions in both eastern Asia and eastern North America<sup>2,3</sup>. However, the temperate deciduous forest biome, which currently harbours many of the eastern Asian/eastern North American disjunct genera we discussed<sup>4</sup>, is poorly represented as a vegetation type in fossil-pollen deposits from the Last Glacial Maximum, casting doubt on the accuracy of the reconstruction by Harrison *et al.*

For example, the palaeovegetation of the Last Glacial Maximum for the vast area between 20° and 30° N and 105° and 120° E was reconstructed from only six pollen

localities, including only one from the interior of eastern Asia, which currently harbours the greatest plant diversity in the region.

It is not possible to determine whether temperate forests coalesced or fragmented during the Last Glacial Maximum without more detailed information. Nonetheless, the reconstruction by Harrison *et al.* confirms that the currently isolated temperate forests of China, Japan, and probably the Korean peninsula, were connected during the Last Glacial Maximum.

Speciation and extinction are processes of populations, not biomes. Climate change causes restructuring of forest communities, probably including the dispersion of temperate deciduous forest taxa among other vegetation types<sup>5</sup>. Vegetation maps cannot provide a detailed picture of past fragmentation and coalescence of particular species populations. Furthermore, because most eastern Asian/eastern North American disjunctions pre-date the onset of major glacial climate cycles in the Northern Hemisphere<sup>6</sup>, diversification of disjunct lineages may have begun under pre-Pleistocene climate conditions that were quite different from those in the present or in the Last Glacial Maximum. In these pre-Pleistocene conditions, eustatic sea-level changes may have been important in isolating and rejoining populations.

Reconstruction of vegetation history, combined with physiographical and climatic heterogeneity, conveys a general impression of the capacity of a region to promote or retard diversification through allopatric speciation<sup>7</sup>. However, the biome reconstruction by Harrison *et al.*, including their interpretation of restricted deciduous forest vegetation during glacial maxima, does not contradict the idea that the more complex landforms and climates of eastern Asia provided greater opportunities for allopatric formation of new species compared with eastern North America.

**Hong Qian\***, **Robert E. Ricklefs†**

\*Research and Collections Center, Illinois State Museum, Springfield, Illinois 62703, USA  
e-mail: hqian@museum.state.il.us

†Department of Biology, University of Missouri-St Louis, St Louis, Missouri 63121, USA

1. Prentice, I. C. et al. *J. Biogeogr.* **19**, 117–134 (1992).
2. Yu, G. et al. *J. Biogeogr.* **27**, 635–664 (2000).
3. Williams, J. W., Webb, T., Richard, P. H. & Newby, P. *J. Biogeogr.* **27**, 585–607 (2000).
4. Qian, H. & Ricklefs, R. E. *Nature* **407**, 180–182 (2000).
5. Webb, T. in *Global Warming and Biological Diversity* (eds Peters, R. L. & Lovejoy, T. E.) 59–75 (Yale Univ. Press, New Haven, Connecticut, 1992).
6. Wen, J. *Annu. Rev. Ecol. Syst.* **30**, 421–455 (1999).
7. Dynesius, M. & Jansson, R. *Proc. Natl Acad. Sci. USA* **97**, 9115–9120 (2000).

**brief communications** is intended to provide a forum for both brief, topical reports of general scientific interest and technical discussion of recently published material of particular interest to non-specialist readers. Priority will be given to contributions that have fewer than 500 words, 10 references and only one figure. Detailed guidelines are available on *Nature's* website ([www.nature.com](http://www.nature.com)) or on request from [nature@nature.com](mailto:nature@nature.com)



# Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1

J. Cliff Yoon\*, Pere Puigserver\*, Guoxun Chen†, Jerry Donovan\*, Zhidan Wu\*, James Rhee\*, Guillaume Adelmant\*, John Stafford‡, C. Ronald Kahn§, Daryl K. Granner‡, Christopher B. Newgard† & Bruce M. Spiegelman\*

\* Dana-Farber Cancer Institute and Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

† Touchstone Center for Diabetes Research, Departments of Biochemistry and Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA

‡ Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA

§ Joslin Diabetes Center and Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA

**Blood glucose levels are maintained by the balance between glucose uptake by peripheral tissues and glucose secretion by the liver. Gluconeogenesis is strongly stimulated during fasting and is aberrantly activated in diabetes mellitus. Here we show that the transcriptional coactivator PGC-1 is strongly induced in liver in fasting mice and in three mouse models of insulin action deficiency: streptozotocin-induced diabetes, *ob/ob* genotype and liver insulin-receptor knockout. PGC-1 is induced synergistically in primary liver cultures by cyclic AMP and glucocorticoids. Adenoviral-mediated expression of PGC-1 in hepatocytes in culture or *in vivo* strongly activates an entire programme of key gluconeogenic enzymes, including phosphoenolpyruvate carboxykinase (PEPCK) and glucose-6-phosphatase, leading to increased glucose output. Full transcriptional activation of the PEPCK promoter requires coactivation of the glucocorticoid receptor and the liver-enriched transcription factor HNF-4 $\alpha$  (hepatic nuclear factor-4 $\alpha$ ) by PGC-1. These results implicate PGC-1 as a key modulator of hepatic gluconeogenesis and as a central target of the insulin–cAMP axis in liver.**

Mammalian cells use glucose as a major energy source; certain cell types such as neurons and red blood cells are especially dependent on glucose. Homeostatic mechanisms are in place to maintain blood glucose levels within a very narrow range, protecting the body against hypoglycaemia during periods of fasting and against excessively high levels following the ingestion of a high-carbohydrate meal. These goals are met chiefly through the hormonal modulation of the production of glucose by the liver and the peripheral uptake of glucose by skeletal muscle, heart muscle and fat.

The liver can produce glucose by breaking down glycogen (glycogenolysis) and by *de novo* synthesis of glucose from non-carbohydrate precursors such as lactate, pyruvate, glycerol and alanine (gluconeogenesis). Glycogenolysis occurs more rapidly, beginning within 2–3 h after a meal in humans, but gluconeogenesis assumes a much greater importance with prolonged fasting<sup>1,2</sup>. The rate of gluconeogenesis is controlled principally by the activities of certain unidirectional enzymes, such as PEPCK, fructose-1,6-bisphosphatase and glucose-6-phosphatase. The genes encoding these proteins are powerfully controlled at the transcriptional level by key hormones, particularly insulin, glucagon and glucocorticoids. In the fasted state and in untreated type 1 diabetes, insulin levels drop while glucagon secretion goes up and catecholamines and glucocorticoids remain stable, resulting in increased glycogenolysis and gluconeogenesis. In the fed state, insulin increases, suppressing glycogenolysis and hepatic glucose production<sup>3</sup>. In both type 1 and type 2 diabetes, excessive hepatic glucose production is a major contributor to both the fasting hyperglycaemia and the exaggerated postprandial hyperglycaemia.

We have previously described a coactivator of nuclear receptors and other transcription factors, PGC-1, which has a major role in cellular respiration and adaptive thermogenesis in tissues such as skeletal muscle and brown fat<sup>4,5</sup>. PGC-1 is induced in muscle and brown fat by exposure of animals to cold; the induction of PGC-1 in brown fat in response to cold is a consequence of cAMP elevation

and is dependent on the  $\beta_3$ -adrenergic receptor<sup>4–7</sup>. These findings provided one of the first reported examples of a transcriptional coactivator linking external stimuli to a specific biological programme of gene regulation. This also raises questions about the role of PGC-1 in tissues where it is expressed but that are not thought to be active in thermogenesis.

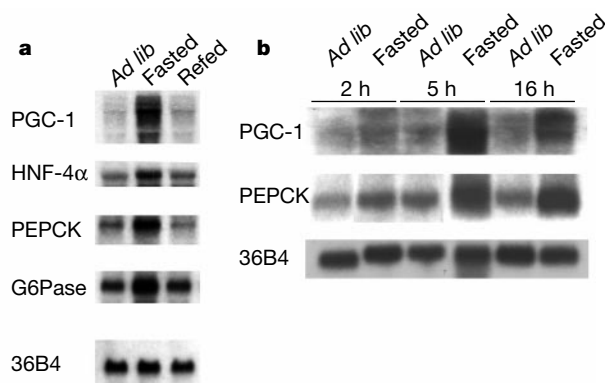
Here, we show that PGC-1 is induced by cAMP and glucocorticoids in isolated hepatocytes and is strongly induced in the liver by fasting and other states of relative insulin deficiency. Furthermore, we show that PGC-1 can powerfully stimulate expression of key genes of the gluconeogenic pathway when expressed in isolated hepatocytes or in the liver of normal rats. These data strongly suggest that regulation of the transcriptional coactivator PGC-1 is of central importance in the control of gluconeogenesis by insulin and other hormones.

## Hormonal modulation of PGC-1 mRNA expression

To study the potential role of PGC-1 in the regulation of liver gene expression by nutritional status, we examined mice subjected to fasting and refeeding. As expected, an overnight fast induced the expression of messenger RNA for the gluconeogenic proteins PEPCK and glucose-6-phosphatase (Fig. 1). HNF-4 $\alpha$ , which is known to be involved in regulation of PEPCK<sup>8</sup>, was also induced. Fasting also induced a parallel 3.7-fold increase in PGC-1 mRNA. All of these inductions were reversed by refeeding (Fig. 1a).

We explored the temporal relationship between the induction of mRNAs for PGC-1 and gluconeogenesis, by performing a time course of fasting (Fig. 1b). An increase in PGC-1 mRNA was first observed after 2 h and peaked 5 h after food deprivation. Induction of PEPCK mRNA was also first detected at 2 h, increasing at 5 and 16 h. Thus, the induction of PGC-1 mRNA is a relatively early event, consistent with a possible role in control of gluconeogenic genes.

Given the known inducibility of PGC-1 in brown fat by  $\beta$ -adrenergic agonists acting through the cAMP system<sup>4–7</sup> and that cAMP is known to rise in the liver during a fast, we investigated



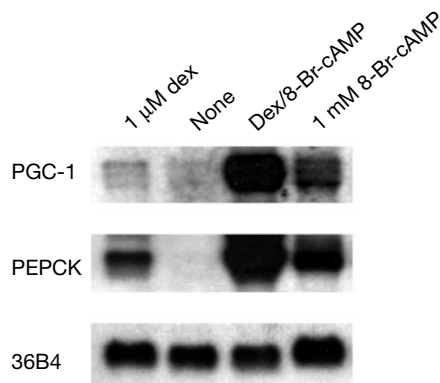
**Figure 1** PGC-1 gene expression is regulated by nutritional status. **a**, Male mice were divided into three experimental groups. The first two groups were either allowed to feed *ad libitum* (lane 1) or subjected to a 24-h fast (lane 2). The third group was subjected to a 24-h fast followed by 24 h of feeding (lane 3), and killed. Ten micrograms of total RNA extracted from pooled liver tissue ( $n = 3$  per group) was analysed by northern blotting. G6Pase, glucose-6-phosphatase catalytic subunit. **b**, Male mice were divided into experimental groups of feeding and fasting (during the night), and killed at the times indicated. RNA liver was extracted and pooled from 2–3 animals per group and subjected to northern analysis.

whether PGC-1 is induced by a cAMP analogue in primary hepatocytes. Cells were treated with 8-bromo-cAMP, dexamethasone or a combination. Whereas dexamethasone produced little increase in the PGC-1 mRNA, treatment with 8-bromo-cAMP caused a large increase in the PGC-1 transcript. Interestingly, the combination of 8-bromo-cAMP and dexamethasone had a synergistic effect (Fig. 2). This is of particular interest as glucocorticoids also rise during the fasted state.

### PGC-1 mRNA is high in insulin-deficient mice

Although gluconeogenic genes are induced in liver by cAMP and glucocorticoids, the dominant regulator of gluconeogenesis *in vivo* is insulin. We first examined the levels of PGC-1 mRNA in the livers of mice treated with streptozotocin, which is the most commonly used experimental model of insulin-deficient type 1 diabetes<sup>9</sup>. PGC-1 mRNA is consistently elevated in the liver of these mice (Fig. 3a).

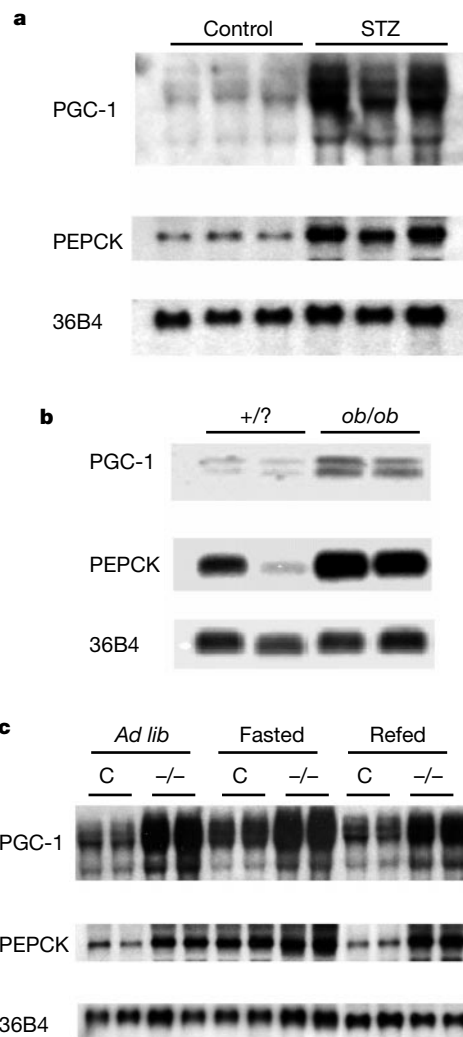
Mice homozygous for the *obesity* (*ob*) gene are obese, severely insulin resistant and widely used as a model of type 2 diabetes.



**Figure 2** PGC-1 gene expression is induced in cultured hepatocytes by treatment with gluconeogenic hormones. Primary hepatocytes were cultured in serum-free DMEM and incubated with vehicle (lane 2), 1  $\mu$ M dexamethasone (dex, lane 1), 1 mM 8-bromo-cAMP (lane 4), or dexamethasone and 8-bromo-cAMP (lane 3). Total RNA was isolated from cells 6–8 h after the initiation of treatment and was subjected to northern analysis (10  $\mu$ g per lane).

PGC-1 mRNA is elevated in the livers of *ob/ob* mice compared with lean controls (Fig. 3b).

The liver-specific insulin-receptor knockout (LIRKO) mouse is a unique model of hepato-specific insulin resistance in the liver, and is associated with impaired glucose tolerance and the failure of insulin to suppress hepatic glucose production<sup>10</sup>. As shown previously, LIRKO mice display increased levels of gluconeogenic gene transcripts in the liver, particularly for PEPCK and glucose-6-phosphatase. A striking elevation of PGC-1 in the livers of the null animals was consistently detected relative to the *lox/lox* controls, correlating well with the expression of gluconeogenic genes in these animals (Fig. 3c). The differences between the knockout and control mice was most apparent in the fed and refed states, although it was still apparent in the fasted state. These data suggest that PGC-1 expression, like gluconeogenesis itself, is suppressed by the action of the insulin receptor.



**Figure 3** PGC-1 gene expression is increased in the livers of insulin-deficient animals. Northern analysis was performed on the total RNA (10  $\mu$ g per lane) isolated from the liver tissues removed from the animals. **a**, Regulation of PGC-1 in streptozotocin (STZ)-diabetic mice. Male mice of 6–8 weeks old received intraperitoneal injections daily with sodium citrate solution (control group) or streptozotocin (100  $\mu$ g per g body mass) for 3 d. After 10 d, the animals ( $n = 3$  per group) were killed. Total RNA was isolated from the liver and analysed. **b**, Regulation of PGC-1 in liver of *ob/ob* mice. Three-month-old male *ob/ob* or lean littermates fed *ad libitum* were killed and RNA was extracted from the liver. **c**, Regulation of PGC-1 in LIRKO mice. Two-month-old female control (*lox/lox*) or LIRKO mice were subjected to *ad libitum* feeding, a 24-h fast, or a 24-h fast and 24 h of refeeding. They were then killed and liver tissue was extracted.

# Expression of PGC-1 increases glucose output

To address the possibility that PGC-1 may regulate aspects of hepatic glucose metabolism, PGC-1 was expressed in primary rat hepatocytes with an adenovirus-based vector. PGC-1 expression at a multiplicity of infection (MOI) of 60 markedly increased the mRNAs that encode several key gluconeogenic enzymes (Fig. 4a). These included PEPCK and glucose-6-phosphatase, which catalyse the first committed and the terminal steps of the gluconeogenic pathway, respectively. Fructose-1,6-bisphosphatase mRNA was also elevated several-fold. Glucocorticoids caused an additional induction of gluconeogenic genes in cells receiving PGC-1.

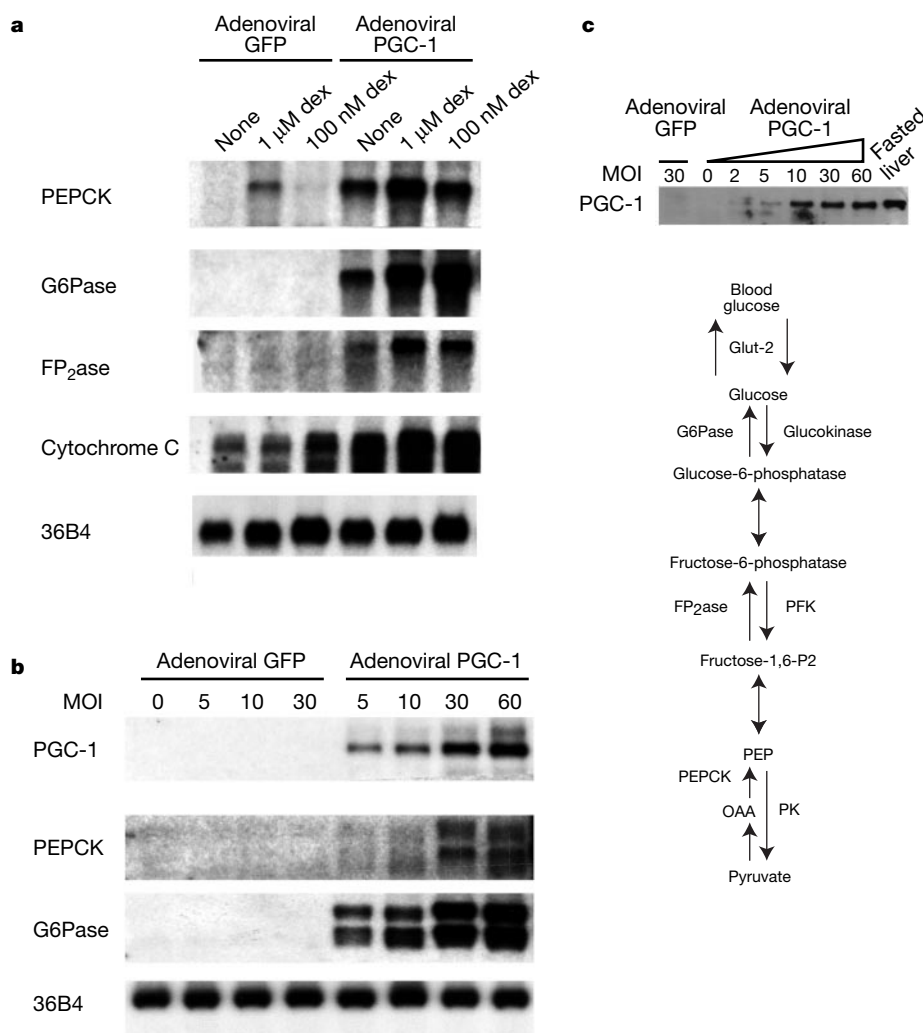
This regulation of gluconeogenic genes is dose-dependent with respect to PGC-1. A significant increase in target gene expression was detected when the hepatocytes were infected at an MOI of 5 (Fig. 4b), and expression of these genes reached a plateau at an MOI of 60. These adenovirus titres resulted in physiological levels of PGC-1 protein: as shown in Fig. 4c, an MOI of 60 resulted in a level of PGC-1 that approaches but does not exceed the level present in

fasted liver. Thus, elevation of PGC-1 stimulated the expression of gluconeogenic genes in the physiological range of this coactivator.

We next examined whether the expression of PGC-1 enhances the net glucose production by primary hepatocytes. Measurements of the glucose released into the culture medium, which initially contained gluconeogenic precursors but no glucose, indicate that the expression of PGC-1 (MOI of 60) by itself is sufficient to increase the glucose production threefold in the absence of any exogenous hormonal stimulation (Fig. 5). The addition of 8-bromo-cAMP did not further enhance this response, which is consistent with the notion that the induction of PGC-1 is the major function of cAMP in this process. Taken together, these data demonstrate that PGC-1, at physiological concentrations, promotes hepatic output of glucose through the transcriptional regulation of the gluconeogenic enzymes.

## PGC-1 coactivates HNF-4 $\alpha$ and the glucocorticoid receptor

The mechanisms underlying the transcriptional activation of the



**Figure 4** Adenovirus-mediated expression of PGC-1 induces multiple genes of the gluconeogenic pathway in hepatocytes. **a**, Regulation of gluconeogenic gene expression by PGC-1. Primary hepatocytes were maintained in DMEM with 10% FBS and at 48 h after plating were infected with control adenovirus expressing GFP or an adenoviral vector expressing both GFP and PGC-1 at an MOI of 60. This titre was sufficient to achieve an infection rate of >90%, as determined by GFP fluorescence. Total RNA was isolated from cells 48 h after infection and analysed by northern blotting (10  $\mu$ g per lane). Dex, dexamethasone; G6Pase, glucose-6-phosphatase; FP<sub>2</sub>ase, fructose-1,6-bisphosphatase. **b**, Dose-dependent activation of gluconeogenic genes by PGC-1. Primary

hepatocytes were infected with increasing titres of GFP or PGC-1 adenovirus and the cells were collected for RNA or protein analysis 48–72 h after infection. Total RNA was probed for the expression of gluconeogenic enzymes, and total cellular protein was immunoblotted for PGC-1 protein expression. **c**, Dose-dependent expression of PGC-1 by adenoviral vectors. PGC-1 was expressed in primary hepatocytes with adenoviral vectors as in **b**. Protein extracts of cells were analysed by immunoblotting for PGC-1 as described in Methods. Fructose-1,6-P2, fructose-1,6-bisphosphate; PFK, phosphofructokinase; PEP, phosphoenolpyruvate; OAA, oxaloacetate; PK, pyruvate kinase.



gluconeogenic enzyme PEPCCK have been extensively studied<sup>2,11</sup>. The activity of this enzyme, which catalyses the first committed step of gluconeogenesis, is regulated primarily at the transcriptional level by a number of hormones, particularly glucagon (via cAMP), glucocorticoids and insulin. Regulatory units, each composed of several DNA-binding elements, have been identified for each of the hormones<sup>8,11–16</sup>. For example, the glucocorticoid response element is composed of transcription factor binding sites (*gAF1*, *gAF2*, *gAF3*) and each of these is necessary for a full hormone-mediated transcriptional response (Fig. 6a; ref. 17). Two relatively weak binding sites exist for the glucocorticoid receptor. However, assigning cAMP and insulin responsiveness of this promoter to the function of a particular transcriptional component has been particularly difficult.

We examined the ability of PGC-1 to activate the gene promoter for PEPCCK in hepatoma cells. The Fao cell line is a well differentiated rat hepatoma cell line that possesses an active gluconeogenic pathway. In the absence of exogenous hormones, PGC-1 robustly activates the wild-type PEPCCK gene promoter, increasing transcription by about tenfold (Fig. 6a). To identify the regulatory elements mediating this effect of PGC-1, mutant constructs were used in which various regulatory elements within the PEPCCK gene promoter were converted to a yeast Gal4 DNA-binding element<sup>18</sup>. As reported previously, mutations in the *gAF* elements do not affect the basal reporter activity. However, a mutation of *gAF1* and/or *gAF3* sites reduced the PGC-1-mediated activation by about 60%. In contrast, a mutation in the *gAF2* or the *CRE* site did not appreciably affect the magnitude of the activation. These results imply that the interactions between PGC-1 and the factors at the *gAF1* and *gAF3* sites are quantitatively important in the regulation of PEPCCK, although other important interactions must also exist.

The *gAF1* and *gAF3* sites are both DR-1 type sequences that can bind several nuclear receptors. Using hepatocyte extracts, the *gAF1* element has previously been shown to bind the proteins HNF-4 $\alpha$  and COUP-TF, both orphan nuclear receptors<sup>8,17</sup>. The *gAF3* element has also been associated with COUP-TF binding<sup>15</sup>. When a Gal4–HNF-4 $\alpha$  construct encoding the Gal4 DNA-binding domain (DBD) fused to the ligand binding or transactivation domain of

the HNF-4 $\alpha$  gene was cotransfected with the *gAF1* or *gAF3* mutant or the *gAF1/gAF3* double-mutant PEPCCK constructs, full PGC-1 activity was restored (Fig. 6b). In contrast, cotransfection of a Gal4–COUP-TF construct produced only a very small effect. Cotransfection of the Gal4 DBD alone had virtually no effect on the PGC-1-mediated activation. Very similar results were obtained at the *gAF3* site (data not shown). These data suggest that the coactivation of HNF-4 $\alpha$  by PGC-1 through the *gAF1* element (and probably the *gAF3* element) constitute a major mechanism of control at this promoter.

To examine these functional interactions at the *gAF1* site with intact transcription factors, we performed reporter gene assays using PGC-1, wild-type HNF-4 $\alpha$  and COUP-TF proteins, and a multimerized *gAF1* response element. To avoid interference with endogenous liver factors, these experiments were done in 3T3 cells. Neither HNF-4 $\alpha$ , COUP-TF nor PGC-1 activates these target sequences alone (Fig. 6c). Although PGC-1 has no ability to coactivate COUP-TF, it dramatically coactivates HNF-4 $\alpha$  without addition of any hormone or HNF-4 $\alpha$  ligand.

Finally, we examined the role of the glucocorticoid response elements in the PEPCCK gene promoter. PGC-1 has previously been shown to bind to and coactivate the glucocorticoid receptor in a ligand-dependent manner<sup>19</sup>. PGC-1 coactivates the PEPCCK promoter through the glucocorticoid receptor also in a ligand-dependent manner (Fig. 6d). This effect is substantially reduced by mutation at the glucocorticoid receptor-1 (GR-1) site, but not the GR-2 site, consistent with previous data indicating a more important regulatory role for the GR-1 site<sup>15</sup>.

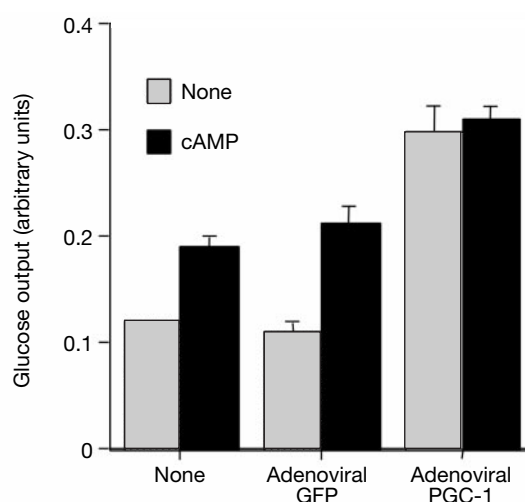
### Interaction with HNF-4 $\alpha$ requires an LXXLL motif

Co-immunoprecipitation experiments were carried out with whole-cell extracts. A Flag-tagged PGC-1 expression construct was transfected into BOSC23 cells with or without HNF-4 $\alpha$ . Immunoprecipitation with anti-Flag antibodies coupled to agarose beads contained significant amounts of the HNF-4 $\alpha$  protein as assessed by western blot analysis, but only when both PGC-1 and HNF-4 $\alpha$  were transfected (Fig. 7a). This indicates that PGC-1 can indeed form a complex with HNF-4 $\alpha$  in cells.

We also performed a series of *in vitro* interaction studies to localize the key domains required for the interaction between PGC-1 and HNF-4 $\alpha$ . Using [<sup>35</sup>S]-labelled *in vitro* translations of HNF-4 $\alpha$  and glutathione S-transferase (GST) fusion proteins of PGC-1, we found that the amino-terminal 190 amino acids of PGC-1 are sufficient to mediate a strong interaction with HNF-4 $\alpha$ , with a recovery of greater than 50% of the input (Fig. 7b). This region of PGC-1 has previously been shown to mediate ligand-dependent interactions with the activation function-2 (AF-2) domains of several nuclear receptors, including oestrogen receptor- $\alpha$  (ER $\alpha$ ), peroxisome proliferator-activated receptor- $\alpha$  (PPAR $\alpha$ ) and glucocorticoid receptor<sup>19–21</sup>. In contrast, PGC-1 interacts with PPAR $\gamma$  in a ligand-independent fashion *in vitro* via PGC-1 amino acids 338–403 (ref. 4).

Because the LXXLL sequence present at amino acids 142–146 of PGC-1 is required for the binding of PGC-1 to other receptors such as ER $\alpha$  and PPAR $\alpha$ , we also tested a mutant construct of PGC-1 (1–190 deletion) in which the LXXLL sequence was mutated. This mutation largely eliminated the binding of PGC-1 to HNF-4 $\alpha$ , identifying this motif as a critical mediator of the physical interaction between PGC-1 and HNF-4 $\alpha$ . To determine whether the loss of ability of the LXXLL mutant to bind HNF-4 $\alpha$  was a specific effect or due to a general loss of proper protein folding, we investigated whether this mutant could still interact with the coactivator SRC-1. The ability to interact with SRC-1 (ref. 22) is unaltered in this mutant (Fig. 7b), suggesting that the PGC-1–HNF-4 $\alpha$  association is indeed mediated specifically by the LXXLL motif.

LXXLL motifs in coactivators of nuclear receptors have been shown to interact with the carboxy-terminal AF-2 domains on the



**Figure 5** Expression of PGC-1 enhances glucose production in hepatocytes in the absence of hormonal treatment. Primary hepatocytes were infected with GFP- or PGC-1-expressing adenovirus and were subsequently cultured in serum-free DMEM with or without 1 mM cAMP treatment. Forty-eight hours after infection, cells were washed twice in phosphate-buffered saline and were incubated in glucose production buffer (see Methods) for 3 h, after which the medium was collected for glucose measurement. Results shown are normalized to the total cell protein content determined from cell lysates.

receptors. A C-terminal-deleted HNF-4 $\alpha$  gene lacking the AF-2 domain expressed HNF-4 $\alpha$  with no ability to bind to PGC-1 (Fig. 7c). Thus, these data strongly suggest that the interaction between PGC-1 and HNF-4 $\alpha$  is mediated by the LXXLL-AF-2 interaction, an interaction that does not require the addition of an exogenous HNF-4 $\alpha$  ligand.

### PGC-1 alters gluconeogenesis *in vivo*

The studies described above were performed with cultured cells; a crucial question is whether elevation of PGC-1 levels could activate gluconeogenesis *in vivo*. We used adenoviral vectors to examine this. Systemic infusion of recombinant adenoviruses into rats through the tail vein has been shown to result primarily in expression of transgenes in the liver, with no detectable expression in peripheral tissues such as muscle, fat, kidney or brain<sup>23,24</sup>. Viruses containing the cDNA encoding PGC-1 (CMV-PGC-1 adenovirus) or green fluorescent protein (CMV-GFP adenovirus) were infused into normal Wistar rats. Five days after virus administration, rats fed *ad libitum* were killed during the day for collection of liver and blood samples. Assays of blood aspartate aminotransferase activity ensured that there was no hepatotoxicity evident in these experiments. Immunoblot analysis revealed that animals who received the CMV-PGC-1 adenovirus had an average increase in PGC-1 protein of 260% relative to animals infused with CMV-GFP. The PGC-1 level in these fed rats was about equal to the levels observed in fasted

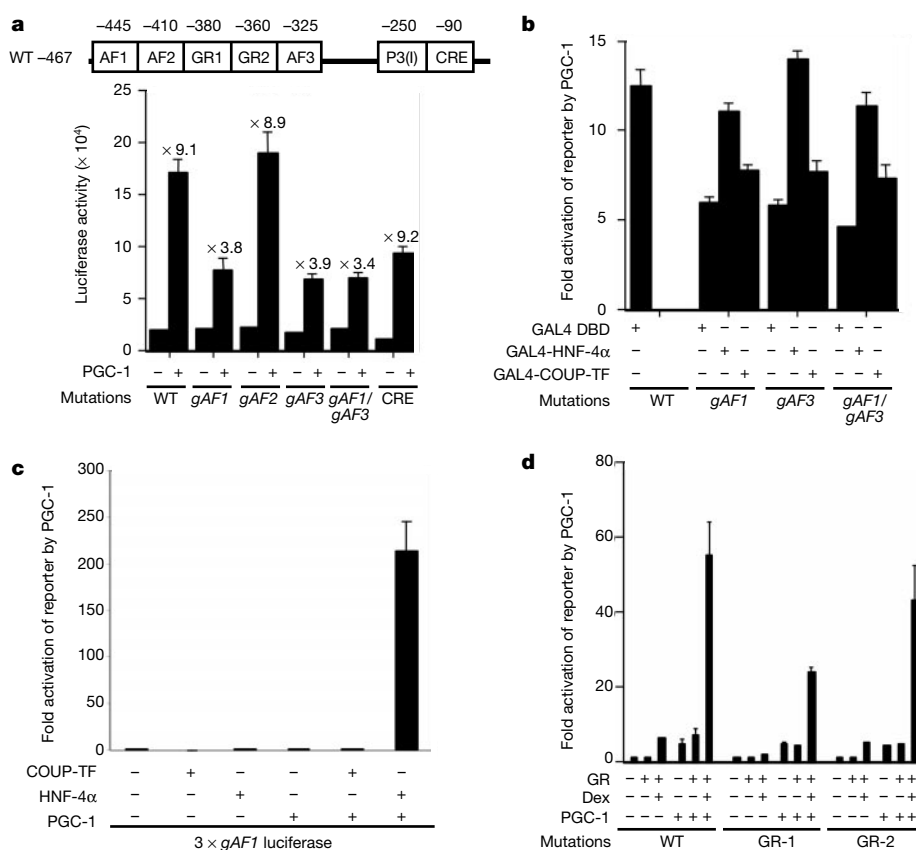
animals (representative blot shown in Fig. 8a).

CMV-GFP-infused control rats had blood glucose levels of  $0.79 \pm 0.06$  mg ml<sup>-1</sup> (mean  $\pm$  s.d.) and insulin levels of  $0.27 \pm 0.03$  ng ml<sup>-1</sup>. In contrast, rats receiving the CMV-PGC-1 adenovirus had glucose levels of  $1.10 \pm 0.03$  mg ml<sup>-1</sup>, a 39% increase ( $P = 0.00028$ ), and  $0.43 \pm 0.05$  ng ml<sup>-1</sup> of insulin, a 66% increase ( $P = 0.0139$ ) (Fig. 8b). Elevated glucose and a compensatory increase of insulin are hallmarks of increased hepatic glucose output in non-diabetic animals; these data closely match what has been observed on overexpression of PEPCK<sup>25</sup> or the catalytic subunit of glucose-6-phosphatase<sup>26</sup> in liver of normal rodents by transgenic and adenoviral methods.

To investigate correlative changes in gene expression in the liver, mRNA levels were examined by northern blotting. The ectopic PGC-1 expression resulted in a dramatic and uniform elevation in mRNA for glucose-6-phosphatase, reaching levels equivalent to those observed in fasted animals (Fig. 8a). There was also an increased expression of PEPCK mRNA, although one control rat also had elevated PEPCK mRNA. These data together show that modulation of PGC-1 levels in the physiological range promotes expression of the gluconeogenesis genes and changes in glucose homeostasis.

### Discussion

Precise regulation of gluconeogenesis is required for the physiological adaptation to fasting and starvation. Carried out mainly in



**Figure 6** PGC-1 activates the PEPCK promoter through a functional interaction with HNF-4 $\alpha$ . **a**, Comparison of the ability of PGC-1 to activate the wild-type (WT) PEPCK gene promoter and various mutant constructs. Fao rat hepatoma cells were maintained in RPMI with 10% FBS and transfected in 6-well plates with 200 ng of pPL32-PEPCK-luciferase (or pGL3-PEPCK-luciferase), 100 ng CMV  $\beta$ -galactosidase and 1  $\mu$ g of pCMV or pCMV-PGC-1 per well. Cells were collected for luciferase assays 48 h after transfection, and the readings were normalized by  $\beta$ -galactosidase activity. **b**, Restoration of the full PGC-1 activity on the gAF1 and gAF3 mutant promoters by cotransfection of HNF-4 $\alpha$ . Vectors encoding the Gal4 DBD, Gal4 DBD-HNF-4 $\alpha$ , or Gal4 DBD-COUP-TF fusion

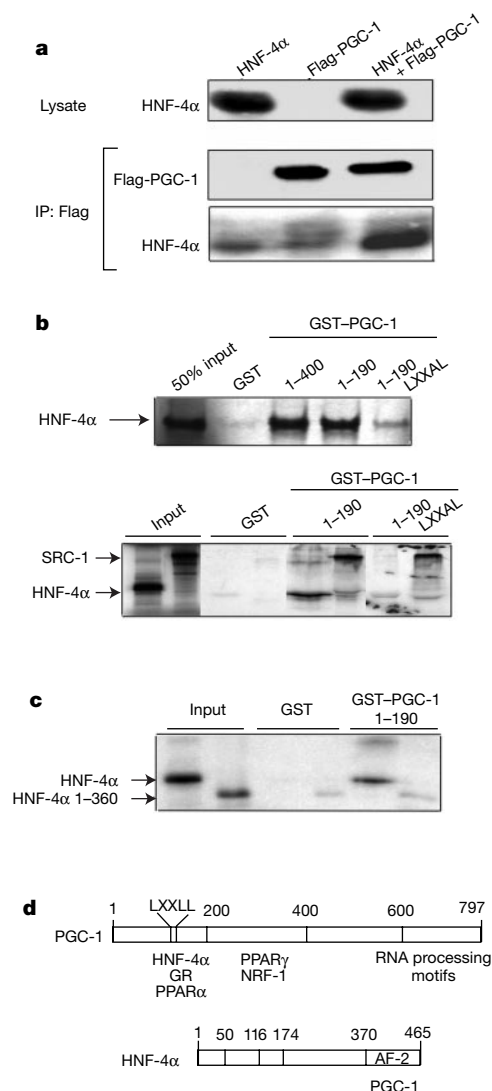
proteins (500 ng per well) were cotransfected with the mutant reporters, containing a substitution of a particular element with the Gal4 DBD, to determine whether they can rescue the PGC-1 activity. **c**, PGC-1 strongly coactivates HNF-4 $\alpha$ , but not COUP-TF on a gAF1 multimerized reporter. NIH 3T3 fibroblasts were transfected as in **a** with vectors encoding proteins as shown. **d**, PGC-1 coactivates the glucocorticoid receptor (GR) in the PEPCK promoter. Fao hepatoma cells were transfected as in **a** with vectors encoding proteins as shown. Dexamethasone (dex, 1  $\mu$ M final concentration) was added for the final 24 h.

the liver, gluconeogenesis enables the organism to maintain relatively steady circulating glucose levels, providing glucose for tissues dependent on this fuel, particularly the brain. Gluconeogenesis is also a particularly important pathological component of type 1 and type 2 diabetes mellitus. The critical role of insulin in suppressing gluconeogenesis and hepatic glucose output is reflected by the fact that excessive hepatic glucose output is a feature of poorly treated diabetes. Increased hepatic glucose output is primarily responsible for fasting hyperglycaemia in diabetes and is a major component of

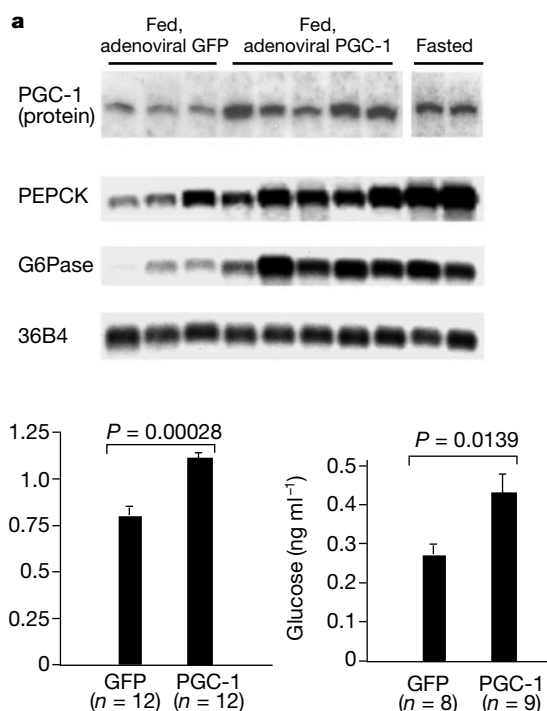
post-prandial hyperglycaemia. Insulin controls hepatic glucose output in significant measure by counteracting the actions of glucagon and catecholamines in the liver. Although these hormones function in gluconeogenesis largely at the transcriptional level, the specific transcription factors targeted by these hormones have been elusive despite considerable experimentation. We show here that a major target of these hormones is the transcriptional coactivator PGC-1.

PGC-1 is induced strongly in primary liver cells by a cAMP analogue and to a lesser extent by a glucocorticoid. However, dexamethasone dramatically increases the cAMP-induced induction of PGC-1. PGC-1 is induced *in vivo* in several states that have been tightly associated with gluconeogenesis—fasting, streptozotocin-induced type 1 diabetes and obesity-linked type 2 diabetes. These all involve a relative insulin deficiency along with an increase in the counter-regulatory hormones glucagon and catecholamines. The dominant effect of insulin in suppressing hepatic gluconeogenesis has been shown definitively by the very marked elevation of PGC-1 that occurs in the LIRKO mouse<sup>10</sup>. PGC-1 is powerfully induced in the livers of these animals. Although this suggests that the PGC-1 induction occurs through the loss of insulin action in the liver *per se*, these animals also show systemic dysfunction<sup>10</sup>. Therefore, indirect actions of insulin cannot be excluded. Recent data have shown important roles for Akt<sup>27</sup> and forkhead transcription factor<sup>28</sup> in insulin suppression of hepatic gluconeogenesis; whether these components are involved in the regulation of PGC-1 gene expression remains to be determined.

That this rise in PGC-1 is causally linked to the induction of gluconeogenesis is illustrated by experiments in which PGC-1 is expressed in primary hepatocytes and rat liver. In hormone-free conditions that minimize endogenous PGC-1 expression in primary hepatocytes, exogenous PGC-1 induced the entire programme of gluconeogenic genes. mRNA for PEPCK, the initial committed



**Figure 7** PGC-1 physically interacts with HNF-4α in cells and *in vitro*. **a**, Co-immunoprecipitation experiments in cells. A Flag-tagged full-length PGC-1 expression construct and a CMV-driven HNF-4α construct were transfected into BOSC23 cells. Whole-cell extracts were made 48 h after transfection and immunoprecipitated with monoclonal antibodies against the Flag tag linked to agarose beads. Immunoprecipitates were analysed by protein immunoblotting with polyclonal antisera against PGC-1 or HNF-4α. **b**, Mapping of the HNF-4α interaction domain within PGC-1. Radiolabelled HNF-4α and SRC-1 proteins were produced by *in vitro* translation with [<sup>35</sup>S]-methionine, and were incubated in a binding buffer with GST control, GST-PGC-1 (amino acids 1–190 and 1–190 with a substitution of Leu 145 to Ala), or GST-PGC-1 (amino acids 1–400) fusion proteins immobilized on glutathione beads. After extensively washing the beads, the bound [<sup>35</sup>S]-labelled HNF-4α and SRC-1 proteins were eluted, separated by SDS-PAGE, and detected by autoradiography. **c**, Mapping of the PGC-1 interaction domain within HNF-4α. Binding interaction of radiolabelled HNF-4α and C-terminus deletion (1–360, generated by polymerase chain reaction) proteins were analysed as in **b**. **d**, Schematic representation of interaction domains in PGC-1 and HNF-4α.



**Figure 8** PGC-1 activates gluconeogenesis *in vivo*. **a**, Male Wistar rats were injected with adenoviruses expressing GFP or PGC-1 as described in the Methods. After 5 d, animals fed *ad libitum* were killed. Livers were collected and analysed for PGC-1 protein and mRNA for PEPCK and glucose-6-phosphatase. Levels of these molecules in livers of fasted, control rats are shown for comparison. **b**, Tail vein blood was analysed for glucose and insulin.



step of this pathway, as well as mRNAs for glucose-6-phosphatase and fructose-1,6-bisphosphatase, were powerfully induced from essentially undetectable levels, in a PGC-1-dose-dependent fashion. The induction of gluconeogenic genes in the primary hepatic cultures resulted in increased glucose secretion. The secretion rate induced by PGC-1 was fully equal to that obtained by cAMP stimulation of cells; cAMP also had no further effect on the glucose secretion rates of cells that contained ectopically expressed PGC-1. Adenoviral-mediated expression of PGC-1 in the liver of rats induced a mild hyperglycaemia and elevation of insulin. This is almost exactly what has been previously observed when hepatic glucose output was activated through adenoviral expression of the gluconeogenic enzymes in rodents<sup>29,30</sup>. Importantly, the changes in expression of gluconeogenic genes were not a consequence of gross overexpression of PGC-1, because the level of ectopic PGC-1 protein obtained in both *in vitro* and *in vivo* studies was equal to that observed in the fasted liver. Taken together, these data indicate that PGC-1 is a major target of the insulin–cAMP axis in inducing hepatic gluconeogenesis.

Because PGC-1 is a transcriptional coactivator, the transcription factor targets of PGC-1 involved in gluconeogenic genes are of considerable importance. PGC-1 powerfully coactivates the PEPCK gene promoter in cotransfection assays and functions prominently through the *gAF1* and *gAF3* sites, apparently through an interaction with HNF-4 $\alpha$ . Reassembly of HNF-4 $\alpha$ , but not of COUP-TF, onto the PEPCK gene promoter completely restores PGC-1 coactivation. PGC-1 could also coactivate HNF-4 $\alpha$  through multimers of its cognate DBDs via direct protein–protein interactions. These data point to a very critical role for HNF-4 $\alpha$  in PGC-1 action at *gAF1*, which also binds COUP-TF and PPAR $\gamma$ . The *gAF3* site has been shown to bind COUP-TF, PPAR $\alpha$  and HNF-4 $\alpha$ . PPAR $\alpha$  is enriched in the liver, is induced in fasting and has been shown previously to be powerfully coactivated by PGC-1. Hence, both PPAR $\alpha$  and HNF-4 $\alpha$  are likely to be targets for PGC-1 at AF-3. It is also clear from these mutation data, particularly the fact that a double *gAF1/gAF3* mutant can still show partial coactivation by PGC-1, that other factors must be important even under these hormone-free conditions. A role for PGC-1 in coactivating the glucocorticoid receptor at glucocorticoid response elements also seems clear.

A function for HNF-4 $\alpha$  in linking PGC-1 to gluconeogenesis is particularly interesting. Gluconeogenesis is a process that takes place primarily in the liver, so the prominent coactivation of the highly liver-enriched HNF-4 $\alpha$  by PGC-1 may serve to tie glucose production to this tissue. From the transcriptional perspective, it is interesting that HNF-4 $\alpha$  has been found previously to be quite active even in the absence of added ligand. Indeed, the nature of the endogenous ligand for this receptor is unclear. However, PGC-1 can powerfully coactivate HNF-4 $\alpha$  in the absence of added ligand. PGC-1 docks on HNF-4 $\alpha$  using the LXXLL motif in the N terminus of PGC-1 and the AF-2 domains of the HNF-4 $\alpha$  gene. Both of these motifs are typically associated with ligand-induced docking on coactivators on nuclear receptors<sup>31</sup>, strongly suggesting that HNF-4 $\alpha$  is either constitutively in a configuration that is similar to what for most receptors is a ligand-induced state, or that an HNF-4 $\alpha$  ligand is nearly ubiquitous in biological fluids. The former seems more likely.

The suppression of gluconeogenesis and hepatic glucose output remains a very attractive therapeutic target in diabetes. The anti-diabetic drug metformin is thought to work through the suppression of hepatic glucose output<sup>32</sup>, but very little is known about its mechanism. Suppressing PGC-1 function in the liver without compromising its effects in other non-gluconeogenic tissues such as brown fat and muscle could yield medically significant anti-diabetic effects. Whereas the development of liver-selective inhibitors of PGC-1 *per se* might be possible, the development of an inhibitor of the PGC-1–HNF-4 $\alpha$  interaction may be more immediately feasible. □

## Methods

### Cell culture

Primary rat hepatocytes were purchased from In Vitro Technologies and cultured in DMEM with 10% FBS at 5% CO<sub>2</sub>. Cells were incubated in serum-free DMEM overnight before addition of hormones. Fao rat hepatoma cells were maintained in RPMI 1640 medium with 10% FBS at 5% CO<sub>2</sub>.

### Animal experiments

Male mice were either allowed free access to food or subjected to a 24-h fast before being killed. A third group was fasted for 24 h and then allowed to feed *ad libitum* for another 24 h before being killed. For the fasting time-course experiment, mice were fed *ad libitum* or fasted during the dark period for 2, 5 and 16 h. For the streptozotocin-diabetic mouse experiments, 6–8-week-old male mice ( $n = 3$  per group) were injected intraperitoneally for 3 d consecutively with either sodium citrate solution or streptozotocin (100  $\mu$ g per g body mass). Animals were killed after 10 d, at which point the mean blood glucose in the streptozotocin-treated group rose to  $>4$  mg ml<sup>-1</sup>. Littermate *ob/ob* and control male mice (Jackson) were killed at 3 months old. For the LIRKO mouse experiments, 2-month-old female control (*lox/lox*) or LIRKO mice were divided into three groups ( $n = 2$  each) and subjected to the feeding–fasting protocol described above.

### Adenovirus infection

Primary hepatocytes were infected 48 h after plating with adenoviruses expressing either GFP or PGC-1 (ref. 33). Cells were collected for RNA or protein isolation 48–72 h after infection. Glucose production assays were performed 48 h after infection.

### Glucose production assay

Primary hepatocytes were cultured in six-well plates ( $1.4 \times 10^6$  cells per well) in DMEM with 10% FBS or, in case of hormonal treatments, in serum-free DMEM. The medium was then replaced with 1 ml of glucose production buffer consisting of glucose-free DMEM (pH 7.4), without phenol red, supplemented with 20 mM sodium lactate and 2 mM sodium pyruvate. After a 3-h incubation, 0.5 ml of medium was collected and the glucose concentration measured with a colorimetric glucose assay kit (Sigma). The readings were then normalized to the total protein content determined from the whole-cell lysates.

### Western blot analysis

Whole-cell extracts were prepared by lysing the cells in a buffer containing 100 mM Tris buffer (pH 8.5), 250 mM NaCl, 1% NP-40, 1 mM EDTA, protease inhibitors (Boehringer Mannheim), and 0.1% phenylmethylsulphonyl fluoride, and were centrifuged at 14,000g for 10 min to remove cellular debris. Tissue extracts were prepared by homogenization in the lysis buffer with a Polytron homogenizer, followed by centrifugation to remove particulate matter. Proteins were separated by SDS–PAGE, transferred to Immobilon P membranes (Millipore), and probed with an affinity-purified antibody directed against murine PGC-1.

### Transcriptional activation assays

NIH 3T3 fibroblasts or FAO hepatoma cells were transiently transfected in six-well dishes at 80–90% by Eugene (Roche). The ratio of DNA:Eugene was 1:2. Culture medium was changed after 24 h. Cells were collected 48 h after transfection and  $\beta$ -galactosidase and luciferase assays were performed. Transfections were performed in duplicate and repeated at least twice.

### Co-immunoprecipitation experiments

Flag-tagged PGC-1 and/or HNF-4 $\alpha$  was transfected into BOSC23 cells using Eugene (Roche). Whole-cell lysates, prepared 48 h after transfection, were subjected to an overnight incubation with a monoclonal antibody to Flag (Sigma) linked to agarose beads<sup>22</sup>. The immunoprecipitates were washed extensively with lysis buffer, separated by SDS–PAGE, and immunoblotted using antibodies directed against PGC-1 and HNF-4 $\alpha$  (Santa Cruz).

### Protein interaction assays

GST fusion proteins produced in *Escherichia coli* were purified on beads containing glutathione. [<sup>35</sup>S]-labelled HNF-4 $\alpha$  was produced with a TNT reticulocyte lysate *in vitro* transcription and translation kit (Promega). About 1  $\mu$ g of fusion protein was mixed with 5  $\mu$ l of the *in vitro* translate in a binding buffer containing 20 mM HEPES buffer (pH 7.7), 75 mM KCl, 0.1 mM EDTA, 2.5 mM MgCl<sub>2</sub>, 0.05% NP40, 2 mM dithiothreitol and 10% glycerol. Binding was performed for 1 h at room temperature and the beads were then washed four times with the binding buffer and resuspended in SDS–PAGE sample buffer. After electrophoresis, the [<sup>35</sup>S]-labelled proteins were detected by autoradiography.

### Adenovirus infusions and metabolic measurements

Male Wistar rats (Charles River) were fed with standard laboratory food and weighed 300–350 g at the time of the studies. Animals were anaesthetized by injection of 0.1 ml per 100 g body mass of a solution containing 25 mg ml<sup>-1</sup> of ketamine (Fort Dodge Animal Health), 2.5 mg ml<sup>-1</sup> of xylazine (Phoenix Scientific) and 0.5 mg ml<sup>-1</sup> of acepromazine (Fermenta Animal Health). CMV-GFP and CMV-PGC-1 adenoviruses were purified by CsCl gradient centrifugation as described<sup>30</sup>. Pure recombinant virus ( $1 \times 10^{12}$

plaque-forming units), suspended in 0.5 ml of PBS, was injected into anaesthetized rats through the tail vein. Animals were allowed to recover and were fed standard food *ad libitum*. Five days after adenovirus infusion, tail vein blood was collected at 14:00 for measurement of glucose concentration, using a  $\beta$ -glucose analyser (HemoCue AB). After the measurement, animals were killed. Blood samples were collected from the heart, centrifuged at 1,300g for 15 min at 4 °C in 15-ml centrifuge tubes containing 50  $\mu$ l of 0.4 M EDTA, and stored at -20 °C before measurement of insulin concentrations with a rat-insulin-specific radioimmunoassay kit (Linco Research). Liver samples were collected, snap frozen in liquid nitrogen, and stored at -80 °C. Aliquots of frozen liver samples were processed for northern blot or immunoblot analysis.

Received 29 June; accepted 1 August 2001.

- Nordlie, R. C. & Foster, J. D. Regulation of glucose production by the liver. *Annu. Rev. Nutr.* **19**, 379–406 (1999).
- Pilkis, S. J. & Granner, D. K. Molecular physiology of the regulation of hepatic gluconeogenesis and glycolysis. *Annu. Rev. Physiol.* **54**, 885–909 (1992).
- DeFronzo, R. A. Pathogenesis of type 2 diabetes: metabolic and molecular implications for identifying diabetes genes. *Diabetes Rev.* **5**, 177–269 (1997).
- Puigserver, P. *et al.* A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* **92**, 829–839 (1998).
- Wu, Z. *et al.* Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**, 115–124 (1999).
- Boss, O. *et al.* Role of the  $\beta_3$ -adrenergic receptor and/or a putative  $\beta_4$ -adrenergic receptor on the expression of uncoupling proteins and peroxisome proliferator-activated receptor- $\gamma$  coactivator-1. *Biochem. Biophys. Res. Commun.* **261**, 870–876 (1999).
- Gomez-Ambrosi, J., Fruhbeck, G. & Alfredo-Martinez, J. Rapid *in vivo* PGC-1 mRNA upregulation in brown adipose tissue of Wistar rats by a  $\beta_3$ -adrenergic agonist and lack of effect of leptin. *Mol. Cell. Endocrinol.* **176**, 85–90 (2001).
- Hall, R. K., Sladek, F. & Granner, D. K. The orphan receptors COUP-TF and HNF-4 serve as accessory factors required for phosphoenolpyruvate carboxykinase gene transcription by glucocorticoids. *Proc. Natl Acad. Sci.* **92**, 412–416 (1995).
- McNeill, J. H. *Experimental Models of Diabetes* (CRC, Boca Raton, 1999).
- Michael, M. D. *et al.* Loss of insulin signaling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. *Mol. Cell* **6**, 87–97 (2000).
- Hanson, R. W. & Reshef, L. Regulation of phosphoenolpyruvate carboxykinase (GTP) gene expression. *Annu. Rev. Biochem.* **66**, 581–611 (1997).
- Mitchell, J. *et al.* Integration of multiple signals through a complex hormone response unit in the phosphoenolpyruvate carboxykinase gene promoter. *Mol. Endocrinol.* **8**, 585–594 (1994).
- Park, E. A., Song, S., Vinson, C. & Roesler, W. J. Role of CCAAT enhancer-binding protein  $\beta$  in the thyroid hormone and cAMP induction of phosphoenolpyruvate carboxykinase gene transcription. *J. Biol. Chem.* **274**, 211–217 (1999).
- Roesler, W. J., Simard, J., Graham, J. G. & McFie, P. J. Characterization of the liver-specific component of the cAMP response unit in the phosphoenolpyruvate carboxykinase (GTP) gene promoter. *J. Biol. Chem.* **269**, 14276–14283 (1994).
- Scott, D. K., Mitchell, J. A. & Granner, D. K. The orphan receptor COUP-TF binds to a third glucocorticoids accessory factor element within the phosphoenolpyruvate carboxykinase gene promoter. *J. Biol. Chem.* **271**, 31909–31904 (1996).
- Yeagley, D., Agati, J. M. & Quinn, P. G. A tripartite array of transcription factor binding sites mediates cAMP induction of phosphoenolpyruvate carboxykinase gene transcription and its inhibition by insulin. *J. Biol. Chem.* **273**, 18743–18750 (1998).
- Sugiyama, T., Wang, J. C., Scott, D. K. & Granner, D. K. Transcriptional activation by the orphan nuclear receptor, chicken ovalbumin upstream promoter-transcription factor I (COUP-TFI). *J. Biol. Chem.* **275**, 3446–3454 (2000).

- Wang, J. C., Stromstedt, P. I., Sugiyama, T. & Granner, D. K. The phosphoenolpyruvate carboxykinase gene glucocorticoid response unit: identification of the functional domains of accessory factors HNF3 $\beta$  (hepatic nuclear factor-3 $\beta$ ) and HNF4 and the necessity of proper alignment of their cognate binding sites. *Mol. Endocrinol.* **13**, 604–618 (1999).
- Knutti, D., Kaul, A. & Kralli, A. A tissue-specific coactivator of steroid receptors, identified in a functional genetic screen. *Mol. Cell. Biol.* **20**, 2411–2422 (2000).
- Tcherepanova, I., Puigserver, P., Norris, J. D., Spiegelman, B. M. & McDonnell, D. P. Modulation of estrogen receptor- $\alpha$  transcriptional activity by the coactivator PGC-1. *J. Biol. Chem.* **275**, 16302–16308 (2000).
- Vega, R. B., Huss, J. M. & Kelly, D. P. The coactivator PGC-1 cooperates with peroxisome proliferator-activated receptor  $\alpha$  in transcriptional control of nuclear genes encoding mitochondrial fatty acid oxidation enzymes. *Mol. Cell. Biol.* **20**, 1868–1876 (2000).
- Puigserver, P. *et al.* Activation of PPAR $\gamma$  coactivator-1 through transcription factor docking. *Science* **286**, 1368–1371 (1999).
- Trinh, K., O'Doherty, R. M., Anderson, P., Lange, A. J. & Newgard, C. B. Perturbations of fuel homeostasis caused by overexpression of the glucose-6-phosphatase catalytic subunit in liver of normal rats. *J. Biol. Chem.* **273**, 31615–31620 (1998).
- O'Doherty, R. M., Lehman, D., Telemaque-Potts, S. & Newgard, C. B. Metabolic impact of glucokinase overexpression in liver: Lowering of blood glucose in fed rats is accompanied by hyperlipidemia. *Diabetes* **48**, 2022–2027 (1999).
- Valera, A., Pujol, A., Pelegrin, M. & Bosch, F. Transgenic mice overexpressing phosphoenolpyruvate carboxykinase gene develop non-insulin-dependent diabetes mellitus. *Proc. Natl Acad. Sci. USA* **91**, 9151–9154 (1994).
- Seoane, J. K. *et al.* Metabolic impact of adenovirus-mediated overexpression of the glucose-6-phosphatase catalytic subunit in hepatocytes. *J. Biol. Chem.* **272**, 26972–26977 (1997).
- Cho, H. *et al.* Insulin resistance and diabetes mellitus-like syndrome in mice lacking the protein kinase Akt2 (PKB $\beta$ ). *Science* **292**, 1728–1731 (2001).
- Nakae, J., Barr, V. & Accili, D. Differential regulation of gene expression by insulin and IGF-1 receptors correlates with phosphorylation of a single amino acid residue in the forkhead transcription factor FKHR. *EMBO J.* **19**, 989–996 (2000).
- Antinozzi, P. A., Berman, H. K., O'Doherty, R. M. & Newgard, C. B. Metabolic engineering with recombinant adenoviruses. *Annu. Rev. Nutr.* **19**, 511–544 (1999).
- Becker, T. *et al.* Use of recombinant adenovirus for metabolic engineering. *Methods Cell Biol.* **43**, 161–189 (1994).
- Glass, C. K. & Rosenfeld, M. D. The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev.* **14**, 121–141 (2000).
- Wiernsperger, N. F. & Bailey, C. J. The antihyperglycaemic effect of metformin: Therapeutic and cellular mechanisms. *Drugs* **58**, 31–39 (1999).
- Lehman, J. J. *et al.* Peroxisome proliferator-activated receptor  $\gamma$  promotes cardiac mitochondrial biogenesis. *J. Clin. Invest.* **106**, 847–856 (2000).

## Acknowledgements

We gratefully acknowledge E. Park for the pGL3 PEPCK-Luc constructs; F. Sladek for the HNF-4 $\alpha$  constructs; V. Yechoor for the tissue samples from the streptozotocin-treated animals; S. Curtis for help with the LIRKO animals; and C.Y. Zhang for help with the fasting experiments. We also thank the members of the Spiegelman laboratory for helpful discussions. P.P. and Z.W. were supported by fellowships from the Juvenile Diabetes Foundation. This work was supported by grants to B.M.S., D.K.G., C.R.K. and C.B.N. from the National Institutes of Health.

Correspondence and requests for materials should be addressed to B.M.S. (e-mail: bruce\_spiegelman@dfci.harvard.edu).

# A high-velocity black hole on a Galactic-halo orbit in the solar neighbourhood

I. F. Mirabel<sup>†‡</sup>, V. Dhawan<sup>\*§</sup>, R. P. Mignani<sup>||</sup>, I. Rodrigues<sup>†</sup> & F. Guglielmetti<sup>†</sup>

<sup>†</sup> Service d'Astrophysique/CEA, CE-Saclay, 91191 Gif/Yvette, France

<sup>‡</sup> Instituto de Astronomía y Física del Espacio, cc5, 1428 Buenos Aires, Argentina

<sup>§</sup> National Radio Astronomy Observatory, PO Box 0, Socorro, New Mexico 87801, USA

<sup>||</sup> European Southern Observatory, Karl-Schwarzschild-Strasse 2, Garching bei Munchen, DE-85740, Germany

<sup>\*</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA

\* These authors contributed equally to this work

Only a few of the dozen or so known stellar-mass black holes have been observed away from the plane of the Galaxy<sup>1</sup>. Those few could have been ejected from the plane as a result of a 'kick' received during a supernova explosion, or they could be remnants of the population of massive stars formed in the early stages of evolution of the Galaxy. Determining their orbital motion should help to distinguish between these options. Here we report the transverse motion (in the plane of the sky) for the black-hole X-ray nova XTE J1118+480 (refs 2–5), from which we derive a large space velocity. This X-ray binary system has an eccentric orbit around the Galactic Centre, like most objects in the halo of the Galaxy, such as ancient stars and globular clusters. The properties of the system suggest that its age is comparable to or greater than the age of the Galactic disk. Only an extraordinary 'kick' from a supernova could have launched the black hole into an orbit like this from a birthplace in the disk of the Galaxy.

The high-galactic-latitude ( $l = 157.7^\circ$ ,  $b = +62.3^\circ$ ) X-ray nova XTE J1118+480 was discovered<sup>2</sup> with the RXTE all-sky monitor on 29 March 2000. It exhibited slow outbursts that lasted about 7 months, with a peak X-ray luminosity of  $4 \times 10^{35} D^2 \text{ kpc}^{-2} \text{ erg s}^{-1}$  and energy spectra typical of black hole binaries in the low/hard state<sup>6</sup>. (Here  $D$  is the distance to the source in units of kpc.) From observations of the  $\sim 19$ -mag optical counterpart<sup>3</sup> in quiescence, a mass function for the compact object  $f(M) \approx 6.0 \pm 0.4 M_\odot$  and an average distance of  $1.85 \pm 0.36 \text{ kpc}$  were determined<sup>4,5</sup>. For approximately 100 days the source exhibited a steady and slowly

variable unresolved radio counterpart<sup>7,8</sup> with a persistent inverted radio spectrum, which was interpreted as optically thick emission from a compact, powerful synchrotron jet<sup>8,9</sup> that could have a size of  $\leq 0.03 \text{ AU}$  (ref. 9).

To measure the transverse motion on the plane of the sky of XTE J1118+480 we carried out observations of the radio counterpart with the Very Long Baseline Array (VLBA) at 15.4 GHz (wavelength  $\lambda = 2 \text{ cm}$ ) and 8.4 GHz ( $\lambda = 3.6 \text{ cm}$ ) on 4 May and 24 July 2000; afterwards it faded below detection. In both epochs the source was unresolved by the synthesized beams of  $1 \times 0.6$  milliarcsec (mas) and  $2 \times 1 \text{ mas}$  respectively, which correspond to physical dimensions smaller than  $0.7 D \text{ kpc}^{-1} \text{ AU}$ , and a brightness temperature greater than  $10^9 \text{ K}$ . Because of the high galactic latitude and low column of interstellar gas along the line of sight ( $N_H \approx 10^{20} \text{ cm}^{-2}$ ; ref. 10), the VLBA images are relatively unaffected by Galactic electron scattering, allowing us to determine in each epoch the position of the compact, unresolved radio source with relative errors of  $\leq 0.35 \text{ mas}$ .

The observations were done in cycles that included the target, the strong calibrator 3C273, a primary calibrator (Cal. 1) used as phase-reference source, and a second extragalactic calibrator (Cal. 2). Their high elevation and good weather ensured that the phase connection was successful. The epoch, source, frequency, bandwidth, position, and flux for the VLBA observations of XTE J1118+480 and the two extragalactic sources used as references for relative astrometry are given in Table 1. The target had inverted and flat spectra, with similar fluxes within 10% of those interpolated from the monitoring with the Ryle<sup>7</sup> and the Very Large Array telescopes (V.D., G. G. Pooley, M. Rupen, R. H. Hjellming, R. N. Ogley and I.F.M.; manuscript in preparation), which confirms that the source was unresolved by the VLBA synthesized beams. Figure 1 shows a drift relative to the extragalactic frame between the two epochs of observation that, after correction for parallax, amounts to a proper motion on the plane of the sky of  $18.3 \pm 1.6 \text{ mas yr}^{-1}$  with position angle  $\text{PA} = 246^\circ \pm 6^\circ$ .

To obtain a further assessment of the proper motion, we have performed an independent measurement in the optical using photographic plates of the Palomar Observatory Sky Survey (POSS I & II), digitized for the Guide Star Catalogue-II project, which cover a time span of about 43 years. By relative astrometry we find that the secondary has a proper motion of  $12.0 \pm 3.2 \text{ mas yr}^{-1}$  with a  $\text{PA} = 240^\circ \pm 15^\circ$ . Such a value is consistent both in magnitude and position angle with the more precise measurement obtained at radio wavelengths (see Fig. 1).

The velocity components  $U$ ,  $V$ , and  $W$  directed to the Galactic centre, rotation direction, and north Galactic pole are derived

**Table 1 VLBA astrometry of XTE J1118+480**

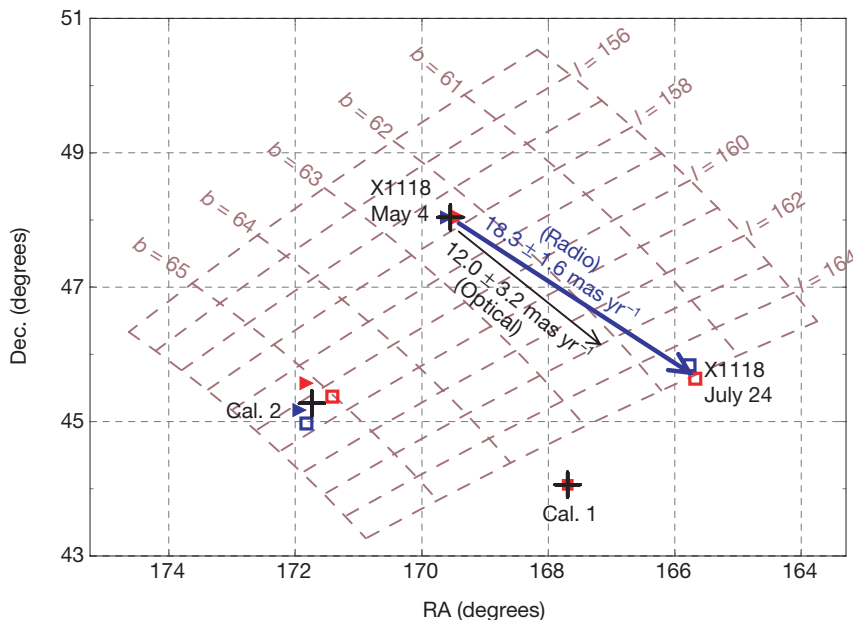
|   | Source | Frequency (GHz) | BW (MHz) | RA (J2000) (h min s ± mas) | Dec (J2000) (° ' " ± mas) | Flux (mJy) |
|---|--------|-----------------|----------|----------------------------|---------------------------|------------|
| A | X1118  | 8.4             | 16       | 11 18 10.7918249 ± 0.093   | 48 02 12.316277 ± 0.131   | 5.9 ± 0.3  |
| A | Cal. 2 | 8.4             | 16       | 11 26 57.6550573 ± 0.014   | 45 16 06.284006 ± 0.023   | 126 ± 3.8  |
| A | Cal. 1 | 8.4             | 16       | 11 10 46.3458105 ± 0.232   | 44 03 25.925163 ± 0.240   | 264 ± 5.3  |
| A | X1118  | 15.4            | 32       | 11 18 10.7918131 ± 0.084   | 48 02 12.316161 ± 0.111   | 8.1 ± 0.3  |
| A | Cal. 2 | 15.4            | 32       | 11 26 57.6550672 ± 0.032   | 45 16 06.283518 ± 0.050   | 68 ± 2.0   |
| A | Cal. 1 | 15.4            | 32       | 11 10 46.3458105 ± 0.232   | 44 03 25.925163 ± 0.240   | 241 ± 4.8  |
| B | X1118  | 8.4             | 64       | 11 18 10.7914440 ± 0.115   | 48 02 12.313919 ± 0.155   | 2.7 ± 0.1  |
| B | Cal. 2 | 8.4             | 64       | 11 26 57.6550224 ± 0.008   | 45 16 06.283783 ± 0.014   | 120 ± 3.2  |
| B | Cal. 1 | 8.4             | 64       | 11 10 46.3458105 ± 0.232   | 44 03 25.925163 ± 0.240   | 267 ± 5.2  |
| B | X1118  | 15.4            | 64       | 11 18 10.7914531 ± 0.159   | 48 02 12.314120 ± 0.191   | 2.9 ± 0.1  |
| B | Cal. 2 | 15.4            | 64       | 11 26 57.6550612 ± 0.007   | 45 16 06.283338 ± 0.014   | 65 ± 1.8   |
| B | Cal. 1 | 15.4            | 64       | 11 10 46.3458105 ± 0.232   | 44 03 25.925163 ± 0.240   | 229 ± 4.6  |

A: MJD 51668, 2000 May 04 UT 03:30–07:40, B: MJD 51749, 2000 July 24 UT 20:00–01:50. Between these two epochs the observed change in position of XTE J1118+480 with respect to the extragalactic frame was  $\Delta \text{RA} = -3.715 \text{ mas}$  and  $\Delta \text{dec.} = -2.149 \text{ mas}$ . The positions at 8.4 and 15.4 GHz were averaged at each epoch and their difference calculated. The position of Cal. 1 with associated absolute error is from the new catalogue of VLBA calibrators (D. Gordon, C. Ma, L. Petrov, A. Beasley, E. Fomalont and A. Peck, personal communication). The errors for other sources are formal fitting errors. The position of Cal. 2 in Table 1 agrees within  $2\sigma$  with its absolute position from the new catalogue. The systematic errors tend to cancel each other out when the difference between the two epochs is taken, and the residual errors in the relative astrometry are smaller than the short-term atmospheric 'seeing' (fluctuations in position caused by residual tropospheric phase errors in the phase-referencing step). The correlator model used for both epochs was the same and contains a seasonally averaged global troposphere, but no short-term weather information. The phase referencing reduces the tropospheric errors to the measured level of 0.3 mas in each axis, as found by breaking the data into 1-h intervals and measuring the root mean square (r.m.s.) position scatter on the source Cal. 2. The errors for XTE J1118+480 are 0.35 mas in each axis, estimated by adding the (maximum) formal error of 0.2 mas in quadrature to the 0.3 mas tropospheric errors.



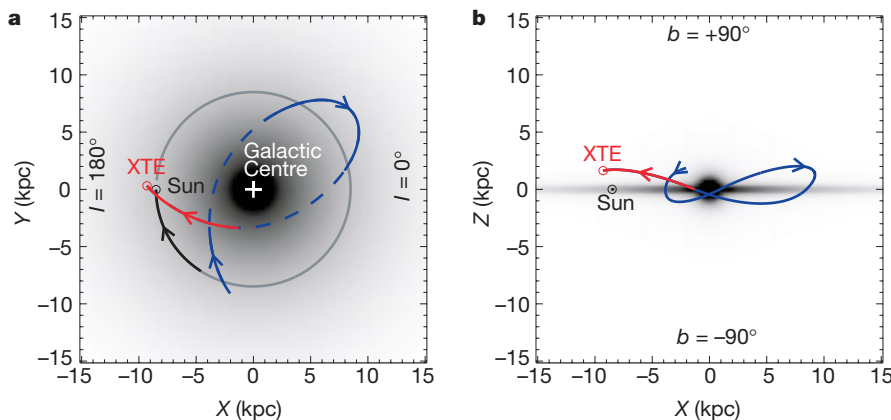
using the equations of transformation<sup>11</sup>, and assuming the Sun moves  $(U_{\odot}, V_{\odot}, W_{\odot}) = (9, 12, 7) \text{ km s}^{-1}$  relative to the local standard of rest (l.s.r.)<sup>12</sup>. For the reasons stated above, in the following we use the proper motion measured with the VLBA which after correction for change in parallax is:  $\Delta \text{RA} = -16.8 \pm 1.6 \text{ mas yr}^{-1}$  and  $\Delta \text{dec.} = -7.4 \pm 1.6 \text{ mas yr}^{-1}$ . Distances of  $1.8 \pm 0.6 \text{ kpc}^4$  and  $1.9 \pm 0.4 \text{ kpc}^5$  are reported for XTE J1118+480, but for the radial velocity of the centre of mass somewhat discrepant

values are estimated<sup>4,5</sup>:  $+26 \pm 17 \text{ km s}^{-1}$  and  $-15 \pm 10 \text{ km s}^{-1}$ , respectively. For a mean value  $d = 1.85 \pm 0.36 \text{ kpc}$  and  $V_r = -15 \pm 10 \text{ km s}^{-1}$ , we find  $(U = -97 \pm 23, V = -101 \pm 23, W = -39 \pm 12) \text{ km s}^{-1}$  in the l.s.r. frame. This implies that the source moves away from the Galactic Centre, has slower rotation about the Galactic Centre than the stars in the disk, and a  $3\sigma$  motion towards the Galactic plane. Adopting  $V_r = +26 \pm 17 \text{ km s}^{-1}$ , we obtain essentially similar results ( $U = -114 \pm 24, V = -94 \pm 23,$



**Figure 1** Proper motion of XTE J1118+480, observed with the VLBA on 4 May–24 July 2000. Shift in position (blue arrow) of the compact radio counterpart of XTE J1118+480 relative to the extragalactic background radio sources Cal. 1 and Cal. 2, measured at  $\lambda = 2 \text{ cm}$  (blue symbols) and  $\lambda = 3.6 \text{ cm}$  (red symbols) on 4 May 2000 (triangles) and 24 July 2000 (squares). The black crosses mark the mean positions. For clarity, the J2000 equatorial coordinates (black lines) and galactic coordinates (dashed brown lines) are shown in expanded scales of  $2^\circ$  and  $1^\circ$  per division respectively. The position of Cal. 1, the primary reference source, is identical at both wavelengths and both epochs, by definition. The position of Cal. 2 is the same at all wavelengths and epochs, within a r.m.s.  $\approx 0.35 \text{ mas}$ . To relate the astrometric positions of XTE J1118+480 in Table 1 to

the galactic frame, a correction for the change in parallax was applied. The position of XTE J1118+480 shifts by  $-16.8 \pm 1.6 \text{ mas yr}^{-1}$  in right ascension and  $-7.4 \pm 1.6 \text{ mas yr}^{-1}$  in declination at both wavelengths. The thin black arrow shows the proper motion of the optical companion measured from a set of four photographic plates that cover a time span of  $\sim 43$  years (from 1953 to 1996). A linear fit to the object positions yielded a proper motion of  $-10.5 \pm 3.2 \text{ mas yr}^{-1}$  in right ascension and of  $-5.9 \pm 3.2 \text{ mas yr}^{-1}$  in declination, after correcting for the galactic rotation and for the peculiar motion of the Sun. The difference in position angle between the radio and optical is  $6^\circ$ , and the errors  $\pm 5^\circ$  and  $\pm 15^\circ$ , respectively.



**Figure 2** Galactic orbit of XTE J1118+480 (blue curve) during the last orbital period of the Sun around the Galactic Centre (240 Myr). The mid-plane mass density distribution of the Galactic bulge and disk are represented by a linear greyscale (darker means dense). The last section of the orbit since the source left the plane  $37.3 \pm 4.8 \text{ Myr}$  ago at a galactocentric distance of  $3.8 \pm 0.5 \text{ kpc}$  is shown in red. The trajectory of the Sun during the later time is indicated by the thick black arc. The source left the plane towards the northern Galactic hemisphere with a galactocentric velocity of  $348 \pm 18 \text{ km s}^{-1}$ , which, after subtraction of the velocity vector due to Galactic rotation, corresponds to a

peculiar space velocity of  $217 \pm 18 \text{ km s}^{-1}$  relative to the Galactic disk frame, and a component perpendicular to the plane of  $126 \pm 18 \text{ km s}^{-1}$ . The orbit of XTE J1118+480 has an eccentricity of 0.54 and is similar to the orbits of halo objects, such as globular clusters. In its oscillating motion (dotted blue line is beneath the Galactic plane) it has just turned around and will fall back in. At the present epoch XTE J1118+480 is at a distance from the Sun of only  $1.85 \pm 0.36 \text{ kpc}$ , flying through the Galactic local neighbourhood with a velocity of  $145 \text{ km s}^{-1}$ . **a**, View from above the Galactic plane; **b**, side view.

$W = -2 \pm 17$  km s<sup>-1</sup>, but with no significant motion perpendicular to the Galactic plane. Within the range of uncertainties for the distance (1.4–2.4 kpc) and radial velocity (–15 to +26 km s<sup>-1</sup>), and irrespective of the specific values adopted, according to the definition of “high-velocity star in the solar neighbourhood”—( $W + 10$ )  $\geq 30$  km s<sup>-1</sup> and/or  $(U^2 + V^2)^{1/2} \geq 65$  km s<sup>-1</sup> (ref. 13)—XTE J1118+480 can be considered a high-velocity X-ray binary, as the velocity relative to the l.s.r. is 145 km s<sup>-1</sup>. The  $V$  component of XTE J1118+480 implies low rotation ( $V_{\phi} \approx 122$  km s<sup>-1</sup>) about the Galactic Centre, contrary to the Galactic disk population, which in the solar neighbourhood is in rapid rotation of around 220 km s<sup>-1</sup>.

Figure 2 shows the orbit of the black hole binary about the Galactic Centre derived from the mean values ( $U = -105 \pm 16$ ,  $V = -98 \pm 16$ ,  $W = -21 \pm 10$ ) km s<sup>-1</sup> in the l.s.r. frame, a distance of  $1.85 \pm 0.36$  kpc from the Sun, and the standard galactic gravitational potential that includes the disk, bulge and halo components<sup>14</sup>. A change of 10% in any of the free parameters of the standard galactic gravitational potential<sup>14</sup> results in changes of less than 5% in any of the orbital parameters of XTE J1118+480.

We now discuss the possibility that XTE J1118+480 could have been launched from the Galactic plane into its halo orbit by a supernova explosion that took place during the formation of the black hole. Using the properties of XTE J1118+480 (refs 4 and 5) (mass of black hole,  $M_{\text{BH}} = 6.9 \pm 0.9 M_{\odot}$ ; mass of binary companion donor star  $M_{\text{donor}} = 0.3 \pm 0.2 M_{\odot}$ ; binary separation of 3 solar radii in circular orbit with orbital period  $P = 0.17$  days), and the equations that describe the impulse from symmetric explosions<sup>15</sup>, it is found that to accelerate the black hole up to a peculiar velocity of 217 km s<sup>-1</sup> by a symmetric explosion more than  $40 M_{\odot}$  would be suddenly ejected during the stellar collapse. This value is implausibly large.

Alternatively, the supernova explosion could be asymmetric and the collapsing star could receive an additional ‘kick’ that can be assumed to be equal for both black holes and neutron stars, the runaway velocity being proportional to the inverse of the mass<sup>16</sup>. Comparing it with the runaway velocities of neutron stars, we find that the momentum of XTE J1118+480 is similar to that of a solitary neutron star moving at 1,000 km s<sup>-1</sup>, about 10 times the average and  $\geq 3$  times the largest linear momentum among millisecond pulsars<sup>17</sup>. Therefore, an origin in the galactic disk would imply that XTE J1118+480 received the most extreme natal impulse among the known binaries that contain compact objects. Because binaries with more massive components are more likely to remain bound after the kicks<sup>16</sup>, and the binary orbit can circularize by tidal friction in a few millions of years (ref. 18), a disk origin through an extraordinary explosion cannot be ruled out.

Instead of being formed in the Galactic disk the black hole in XTE J1118+480 may be the relic of an ancient massive star formed in the Galactic halo. The values of  $U$  and  $V$  are consistent with the large random motions of old halo stars flying through the solar neighbourhood<sup>19</sup> that have metallicities  $[\text{Fe}/\text{H}] = -1.2$  to  $-1.4$ . Furthermore, the Galactic orbit shown in Fig. 2 is similar to that of some globular clusters (for example, NGC6656 with  $[\text{Fe}/\text{H}] = -1.7$ ) (ref. 20). We point out that there are two additional observations that are consistent with the hypothesis of a halo origin: (1) the very low metallicity  $Z/Z_{\odot} \approx 0.1$  of the reflector medium derived from broad-band X-ray spectroscopy<sup>21</sup>, and (2) the depletion of carbon and enhancement of nitrogen found in the ultraviolet spectrum with the Hubble Space Telescope (HST), which would require a long-lasting carbon–nitrogen–oxygen cycle, suggesting that the secondary has lost a large fraction of its mass and is presently exposing the layers that were originally below the surface<sup>22</sup>. Therefore, XTE J1118+480 may be contemporary with globular clusters, being one of the  $10^4$ – $10^5$  black holes that were ejected from these old stellar systems<sup>23</sup> and at present swirl around in the halo.

In the year 2000 XTE J1118+480 brightened from quiescence by

around 6 mag. Our analysis of the historical optical plates reveals that in 1995 and 1996 it had optical outbursts of  $\sim 2$  mag that passed without notice, although they could easily have been measured. We point out that XTE J1118+480 was a black-hole nova with a remarkably large optical-to-X-ray flux ratio<sup>10</sup> that reached a peak luminosity of about  $10^{36}$  erg s<sup>-1</sup> in the 1–160 keV band<sup>6</sup>. At present it is close to the Sun, but at the distance of the Galactic Centre it would have not been detected in most surveys using X-ray instruments with a large field of view. Therefore, it is possible that many faint X-ray binaries have been missed, which, like XTE J1118+480, may sporadically appear as microquasars<sup>24</sup>; that is, as sources of collimated beams of relativistic particles and high energy photons. The association of these microquasars with a subset of unidentified  $\gamma$ -ray sources<sup>25</sup> that are rather variable, soft, faint, and have—as do globular clusters—a scale-height above the galactic plane of  $\sim 2$  kpc, is an intriguing possibility. □

Received 26 March; accepted 31 July 2001.

- White, N. E. & van Paradijs, J. The Galactic distribution of black hole candidates in low-mass X-ray binary systems. *Astrophys. J.* **473**, L25–L29 (1996).
- Remillard, R., Morgan, E., Smith, D. & Smith, E. XTE J1118+480. *IAU Circ. No.* 7389 (2000).
- Uemura, M. *et al.* Discovery and photometric observation of the optical counterpart in a possible Galactic halo X-ray transient, XTE J1118+480. *Publ. Astron. Soc. Jpn* **52**, L15–L19 (2000).
- McClintock, J. E. *et al.* A black hole of greater than  $6 M_{\odot}$  in the X-ray nova XTE J1118+480. *Astrophys. J.* **551**, L147–L150 (2001).
- Wagner, R. M. *et al.* The halo black hole X-ray transient XTE J1118+480. *Astrophys. J.* **556**, 42–46 (2001).
- McClintock, J. E. *et al.* Complete simultaneous spectral observations of the black-hole X-ray nova XTE J1118+480. *Astrophys. J.* **555**, 477–482 (2001).
- Pooley, G. G. & Waldram, E. M. XTE J1118+480. *IAU Circ. No.* 7390 (2000).
- Dhawan, V., Pooley, G. G., Ogley, R. N. & Mirabel, I. F. XTE J1118+480. *IAU Circ. No.* 7395 (2000).
- Fender, R. P. *et al.* Spectral evidence for a powerful compact jet from XTE J1118+480. *Mon. Not. R. Astron. Soc.* **322**, L23–L27 (2001).
- Hynes, R. L. *et al.* The X-ray transient XTE J1118+480: Multiwavelength observations of a low-state minioutburst. *Astrophys. J.* **539**, L37–L40 (2000).
- Johnson, D. R. H. & Soderblom, D. R. Calculating Galactic space velocities and their uncertainties. *Astron. J.* **93**, 864–867 (1987).
- Mihalas, D. & Binney, J. *Galactic Astronomy* 382–384 (Freeman, San Francisco, 1981).
- Oort, J. H. Asymmetry in the distribution of stellar velocities. *Observatory* **49**, 302–304 (1926).
- Johnston, K. V., Hernquist, L. & Bolte, M. Fossil signatures of ancient accretion events in the halo. *Astrophys. J.* **465**, 278–287 (1996).
- Nedmans, G., Tauris, T. M. & van den Heuvel, E. P. J. Constraints on mass ejection in black hole formation derived from black hole X-ray binaries. *Astron. Astrophys.* **352**, L87–L90 (1999).
- Fryer, C. L. & Kalogera, V. Theoretical black hole mass distributions. *Astrophys. J.* **554**, 548–560 (2001).
- Toscano, M. *et al.* Millisecond pulsar velocities. *Mon. Not. R. Astron. Soc.* **307**, 925–933 (1999).
- Zahn, J.-P. Tidal friction in close binary stars. *Astron. Astrophys.* **57**, 383–394 (1977).
- Chiba, M. & Beers, T. C. Kinematics of metal-poor stars in the Galaxy. III. *Astron. J.* **119**, 2843–2865 (2000).
- Dauphine, B. *et al.* The kinematics of globular clusters, apocentric distances and a halo metallicity gradient. *Astron. Astrophys.* **313**, 119–128 (1996).
- Frontera, F. *et al.* A measurement of the broad-band spectrum of XTE J1118+480 with BeppoSAX and its astrophysical implications. *Astrophys. J.* (in the press).
- Haswell, C. A. *et al.* XTE J1118+480. *IAU Circ. No.* 7407 (2000).
- Kulkarni, S. R., Hut, P. & McMillan, S. Stellar black holes in globular clusters. *Nature* **364**, 421–423 (1993).
- Mirabel, I. F. & Rodríguez, L. F. Microquasars in our Galaxy. *Nature* **392**, 673–676 (1998).
- Grenier, I. A. in *GeV–TeV Gamma Ray Astrophysics Workshop: Towards A Major Atmospheric Cherenkov Detector VI* (eds Dingus, B. L., Salamon, M. H. & Kieda, D. B.) Vol. 515, 261–279 (American Institute of Physics, Melville, New York, 1999).

## Acknowledgements

We thank G. Pooley for information on the fluxes observed with the Ryle telescope, A. Spagna for providing his astrometric software for the optical measurements, and R. Fernández for help with software programming. I.F.M. thanks J. Paul, J.-P. Zahn, R. M. Wagner, G. Israelian, R. Rebolo, E. Ergma and A. King for helpful comments and discussions. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The Guide Star Catalogue-II is a joint project of the Space Telescope Science Institute and the Osservatorio Astronomico di Torino with additional support provided by the European Southern Observatory, Space Telescope European Coordinating Facility, the International GEMINI project and the European Space Agency Astrophysics Division. I.F.M. is a member of the Consejo Nacional de Investigaciones Científicas y Técnicas of Argentina, and I.R. is a Fellow of the Conselho Nacional de Desenvolvimento Científico e Tecnológico of Brazil.

Correspondence and requests for materials should be addressed to I.F.M. (e-mail: fmirabel@cea.fr).

# Observation of the Kapitza–Dirac effect

Daniel L. Freimund\*, Kayvan Aflatooni\*† & Herman Batelaan\*

\* Department of Physics and Astronomy, University of Nebraska—Lincoln, 116 Brace Laboratory, PO Box 880111, Lincoln, Nebraska 68588-0111, USA

In their famous 1927 experiment, Davisson and Germer observed<sup>1</sup> the diffraction of electrons by a periodic material structure, so showing that electrons can behave like waves. Shortly afterwards, Kapitza<sup>2</sup> and Dirac<sup>3</sup> predicted that electrons should also be diffracted by a standing light wave<sup>4</sup>. This Kapitza–Dirac effect is analogous to the diffraction of light by a grating, but with the roles of the wave and matter reversed. The electron and the light grating interact extremely weakly, via the ‘ponderomotive potential’<sup>5</sup>, so attempts to measure the Kapitza–Dirac effect had to wait for the development of the laser. The idea<sup>6</sup> that the underlying interaction with light is resonantly enhanced for electrons in an atom led to the observation<sup>7</sup> that atoms could be diffracted by a standing wave of light. Deflection of electrons by high-intensity laser light, which is also a consequence of the Kapitza–Dirac effect, has also been demonstrated<sup>8</sup>. But the coherent interference that characterizes wave diffraction has not hitherto been observed<sup>9,10</sup>. Here we report the diffraction of free electrons from a standing light wave—a realization of the Kapitza–Dirac effect as originally proposed.

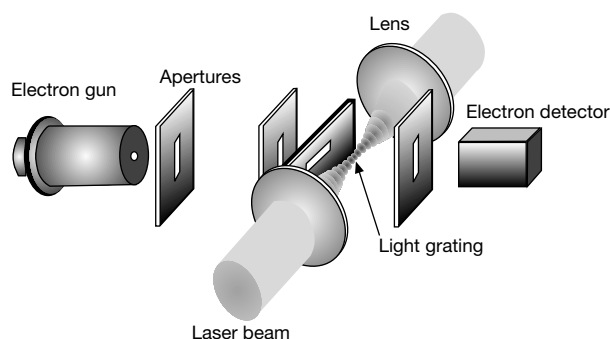
In our experiment, an electron beam crosses two counter-propagating laser beams which form the standing wave light grating (Fig. 1). To reach sufficiently high laser intensities, we used a Nd:YAG laser with 10-ns pulses and an energy of 0.2 J per pulse focused to a beam waist 125  $\mu\text{m}$  in diameter. Each counter-propagating laser beam travels an equal distance not differing by more than 1 mm. This is well within the coherence length of the laser beam (5 mm) where the standing wave is formed. A 380-eV electron beam is collimated by two 10- $\mu\text{m}$ -wide molybdenum slits separated by 24 cm. A third slit cuts the height of the electron beam to the size of the laser beam waist. Subsequently, the electron beam crosses the standing wave about 1 cm after the third slit. A fourth 10- $\mu\text{m}$  slit, 24 cm downstream from the interaction region, is used to scan the electron beam profile. The measured spatial width (full-width at half-maximum, FWHM) of the electron beam is 25  $\mu\text{m}$ . This is a considerably narrower width than the expected distance between the zero and first diffraction order,  $55 \mu\text{m} = 2\lambda_{\text{dB}}/\lambda_{\text{opt}} (\times 24 \text{ cm})$ , where  $\lambda_{\text{dB}}$  is the de Broglie wavelength of the electrons and  $\lambda_{\text{opt}}$  is the wavelength of the laser light, 532 nm. We may thus expect the diffraction peaks to be resolved. The factor of two takes into account the ratio between the light grating periodicity and the light wavelength. The electrons are detected as a function of time with an electron multiplier. Each laser pulse is used as a start signal, and the detection of electrons is used as the stop signal for a time to amplitude converter. A multi-channel scaler records the pulses from the converter into coincidence time spectra. From the time spectra taken at various positions, the diffraction pattern is obtained directly.

The diffraction pattern is shown in Fig. 2. The diffraction orders are clearly resolved and fall at their expected positions ( $n \times 55 \mu\text{m}$ ,  $n = 0, \pm 1, \pm 2, \dots$ ). The heights of the diffraction peaks might be expected to be given by the analytic solution of the Schrödinger equation in the diffractive limit<sup>11</sup>. However, this is not the case. Given that some electrons pass through less intense regions of the

focused laser beam and some electrons pass through more intense regions, a numerical solution of the Schrödinger equation gives acceptable agreement with the experimental data (Fig. 2). The parameters used in the numerical simulation (laser focus, 125  $\mu\text{m}$ ; laser intensity in the standing wave,  $5 \times 10^{14} \text{ W m}^{-2}$ ; electron velocity,  $1.1 \times 10^7 \text{ m s}^{-1}$ ; optics transmission, 70%; overlap, 45  $\mu\text{m}$ ) are consistent with the experimental parameters. An overlap of 45  $\mu\text{m}$  indicates the FWHM of the height of the standing wave. We calculate that with perfect overlap (standing wave FWHM of 125  $\mu\text{m}$ ) between the two counter-propagating laser beams, a laser light intensity ten times lower would yield a comparable diffraction pattern. The small asymmetry in the diffraction pattern (somewhat larger in the experiment than in the simulation) is attributed to a misalignment of the electron beam of approximately 1 mrad with respect to the laser and is indicative of the onset of Bragg scattering.

In some early experiments<sup>12–15</sup> attempts were made to measure the deflection of free electrons due to a light wave. Two experiments reported an effect<sup>12,13</sup>, while two others did not<sup>14,15</sup>. Regardless of this controversy no diffraction peaks were observed. Indeed, recent reviews state that the Kapitza–Dirac effect has not been observed for electrons<sup>9,10</sup>. Explanations were offered to account for the controversy of the early experiments. Schwartz<sup>16</sup> has suggested that in two experiments the interaction strength was accidentally such that the height of the first-order diffraction peak was at a minimum. Considering the experimental difficulty of obtaining uniform laser intensity, this explanation seems unlikely. Fedorov<sup>17</sup>, on the other hand, has suggested that a slow adiabatic turn-on is the main reason for the previous failure to observe the deflection owing to the ‘ponderomotive potential’. In agreement with Fedorov, our simulation also shows that increasing the laser beam spatial width causes the Kapitza–Dirac effect to vanish for finite-sized electron beams. We have kept Fedorov’s suggestion in mind while designing this experiment. Additionally, the greater stability and reliability of modern lasers and the improved performance of electronics have aided this experiment compared to earlier attempts to observe the Kapitza–Dirac effect.

Our results demonstrate that no fundamental problems stood in the way of observing the effect. At much higher laser intensities the important 1988 experiment<sup>8</sup> by Bucksbaum *et al.* showed that electrons could be deflected by the ponderomotive potential. Bucksbaum observed two classical rainbow scattering peaks separated by about 1,000 photon recoils. We observe quantum mechanical diffraction peaks separated by two photon recoils. An important difference between these experiments is that the rainbow peaks are not coherent, whereas diffraction peaks are coherent.



**Figure 1** Schematic of our apparatus. Electrons are collimated by four molybdenum slits and diffract from a standing wave of light formed by two counter-propagating laser beams. The electrons must be described by a quantum mechanical wave while the standing light wave acts as a grating.

† Present address: Department of Physics, Fort Hayes State University, 600 Park Street, Hays, Kansas 67601-4099, USA.



The observation of the Kapitza–Dirac effect opens the door to various new experiments. Because the diffracted electron beams are coherent with each other, the Kapitza–Dirac effect constitutes a coherent beam splitter. Just as for atoms, the combination of three such beam splitters can be used to construct a Mach–Zehnder interferometer<sup>18</sup>. Compared to biprism electron interferometers, this new type of electron interferometer would operate at very low electron energies and seems to be well suited to study, for example, forward electron–atom scattering phase shifts<sup>19</sup>. Instead of using three consecutive beam splitters, it may also be possible to use the coherence of the diffraction pattern itself. When I<sub>2</sub> molecules are placed in a YAG laser beam (with experimental parameters almost identical to those used in our experiment) they will be aligned along the laser polarization axis<sup>20,21</sup>, but only at the antinodes of the standing wave. The result is that the periodically aligned I<sub>2</sub> molecules will write a sinusoidal phase shift on the incoming electron waves. This shift will modify the diffraction pattern and could be used to monitor the I<sub>2</sub> alignment as it is influenced by, for example, molecular dissociation or ionization.

Apart from the use of the Kapitza–Dirac effect as a tool, it is interesting to study in itself. It has been shown experimentally that atoms moving through a standing light wave represent an example of classical and quantum chaos. The largest angles to which atoms can be deflected are determined by the boundary between regular and chaotic motion<sup>22</sup>, and shaking the standing wave back and forth leads to the observation of Anderson localization<sup>23</sup>. Our experiment shows that the same experimental regime can be reached for electrons. The charge of the electron affords a convenient means of studying the effect of external interactions on quantum chaotic behaviour.

Increasing the laser intensity to 10<sup>15</sup> W cm<sup>−2</sup> (which is readily

achieved in 100-ps pulse Nd:YAG lasers<sup>24</sup>) will raise the strength of the magnetic field of the laser beam to the extent that the electron spin would rotate by 180° in such a field. The question thus arises of whether the electron spin in the diffraction process could flip. Although classical arguments for a circularly polarized travelling wave seem to rule out this possibility<sup>24</sup>, this question, in general, and in particular for standing waves, is to our knowledge unanswered. The atom optics counterpart of this effect is the “optical Stern–Gerlach effect” and has been observed<sup>25</sup>. However, this result cannot easily be extended to free electrons owing to the half-integer value of the spin. A spin flip in combination with diffraction would constitute a polarizing beam splitter for free electrons or, in other words, a microscopic Stern–Gerlach magnet. We have to keep in mind that Stern–Gerlach magnets for free electrons do not exist<sup>26</sup>. By increasing the laser intensity further to 10<sup>18</sup> W cm<sup>−2</sup> (for a laser wavelength of 1 μm), it is interesting to note that electrons are so light that relativistic speeds can be reached<sup>24</sup>. Thus the study of the interaction of free electrons with laser light can probably be extended from quantum mechanics to include spin, chaotic behaviour and relativistic mechanics. □

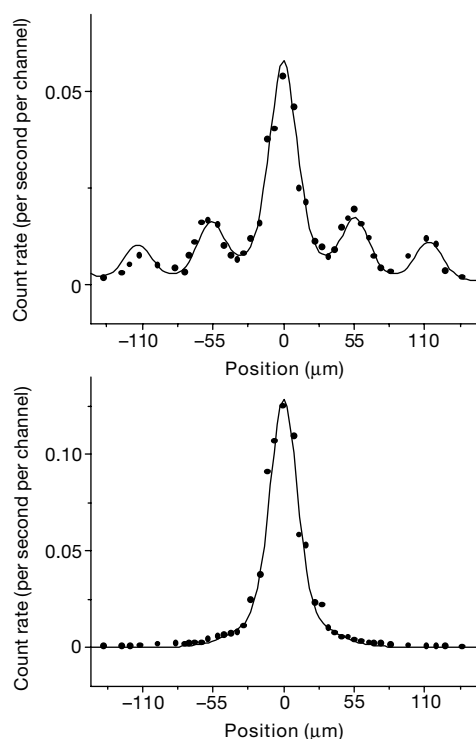
Received 18 May; accepted 19 July 2001.

1. Davisson, C. & Germer, L. H. The scattering of electrons by a single crystal of nickel. *Nature* **119**, 558–560 (1927).
2. Mott, N. Man of courage. *Nature* **350**, 31 (1991).
3. Aitchison, I. Spin doctors. *Nature* **392**, 771–773 (1998).
4. Kapitza, P. L. & Dirac, P. A. M. The reflection of electrons from standing light waves. *Proc. Camb. Phil. Soc.* **29**, 297–300 (1933).
5. Boot, H. A. H. & Harvie, R. B. R.-S. Charged particles in a non-uniform radio-frequency field. *Nature* **180**, 1187 (1957).
6. Althuler, S., Frantz, L. M. & Brauneis, R. Reflection of atoms from standing light waves. *Phys. Rev. Lett.* **17**, 231–232 (1966).
7. Gould, P. L., Ruff, G. A. & Pritchard, D. E. Diffraction of atoms by light: the near-resonant Kapitza–Dirac effect. *Phys. Rev. Lett.* **56**, 827–830 (1986).
8. Bucksbaum, P. H., Schumacher, D. W. & Bashkansky, M. High intensity Kapitza–Dirac effect. *Phys. Rev. Lett.* **61**, 1182–1185 (1988).
9. Adams, C. S., Sigel, M. & Mlynek, J. Atom optics. *Phys. Rep.* **240**, 143–210 (1994).
10. Fedorov, M. V. in *Laser Science and Technology; An International Handbook* No. 13, 1–77 (Harwood Academic, New York, 1991).
11. Batelaan, H. The Kapitza–Dirac effect. *Contemp. Phys.* **41**, 369–381 (2000).
12. Bartell, L. S., Roskos, R. R. & Thompson, H. B. Reflection of electrons by standing light waves: experimental study. *Phys. Rev.* **166**, 1494–1504 (1968).
13. Schwartz, H., Tourtelotte, H. A. & Gaertner, W. W. Direct observation of nonlinear scattering of electrons by laser beam. *Phys. Lett.* **19**, 202–203 (1965).
14. Takeda, Y. & Matsui, I. Electron reflection by standing wave of giant laser pulse. *J. Phys. Soc. Jpn* **25**, 1202 (1968).
15. Pfeiffer, H.-Chr. Experimentelle prüfung der streuwahrscheinlichkeit für elektronen beim Kapitza–Dirac-effekt. *Phys. Lett. A* **26**, 362–363 (1968).
16. Schwarz, H. The Kapitza–Dirac effect at high laser intensities. *Phys. Lett. A* **43**, 457–478 (1973).
17. Fedorov, M. V. Stimulated scattering of electrons by photons and adiabatic switching on hypothesis. *Opt. Commun.* **12**, 205–209 (1974).
18. Rasel, E., Oberthaler, M. K., Batelaan, H., Schmeidmayer, J. & Zeilinger, A. Atom wave interferometry with diffraction gratings of light. *Phys. Rev. Lett.* **75**, 2633–2637 (1995).
19. Forrey, R. C., Dalgarno, A. & Schmeidmayer, J. Determining the electron forward-scattering amplitude using electron interferometry. *Phys. Rev. A* **59**, R942–R945 (1999).
20. Meshulach, D. & Silverberg, Y. Coherent quantum control of two-photon transitions by a femto-second laser pulse. *Nature* **396**, 239–242 (1998).
21. Larsen, J. J., Wendt-Larsen, I. & Stapelfeldt, H. Controlling the branching ratio of photodissociation using aligned molecules. *Phys. Rev. Lett.* **83**, 1123–1126 (1999).
22. Robinson, J. C. *et al.* Can a single-pulse standing wave induce chaos in atomic motion? *Phys. Rev. Lett.* **76**, 3304–3307 (1996).
23. Steck, D. A., Milner, V., Oskay, W. H. & Raizen, M. G. Quantitative study of amplitude noise effects on dynamical localization. *Phys. Rev. E* **62**, 3461–3475 (2000).
24. Bucksbaum, P. H. *Atoms in Strong Fields* (ed. Nicolaides, C. A.) 381–405 (Plenum, New York, 1990).
25. Sleator, T., Pfau, T., Balykin, V., Carnal, O. & Mlynek, J. Experimental demonstration of the optical Stern–Gerlach effect. *Phys. Rev. Lett.* **68**, 1996–1999 (1992).
26. Batelaan, H., Gay, T. J. & Schwendiman, J. J. Stern–Gerlach effect for electron beams. *Phys. Rev. Lett.* **79**, 4517–4521 (1997).

# Acknowledgements

We thank P. Burrow, G. Gallup and T. Gay for discussions. This work was supported by the Research Corporation, the NRI and the NSF Experimental Program to Stimulate Competitive Research.

Correspondence and requests for materials should be addressed to H.B. (e-mail: hbatelaan2@unl.edu).



**Figure 2** Experimental data. The electron detection rate is presented as a function of detector position. Our data (black points) agree reasonably well with a numerical solution of the Schrödinger equation (described in the text) and clearly show diffraction peaks, which is the signature of the Kapitza–Dirac effect. The bottom figure shows the electron beam profile with the laser beams turned off.

# Unexpected inhibition of fusion in nucleus–nucleus collisions

A. C. Berriman, D. J. Hinde, M. Dasgupta, C. R. Morton, R. D. Butt & J. O. Newton

Department of Nuclear Physics, Research School of Physical Sciences and Engineering, Australian National University, Canberra, ACT 0200, Australia

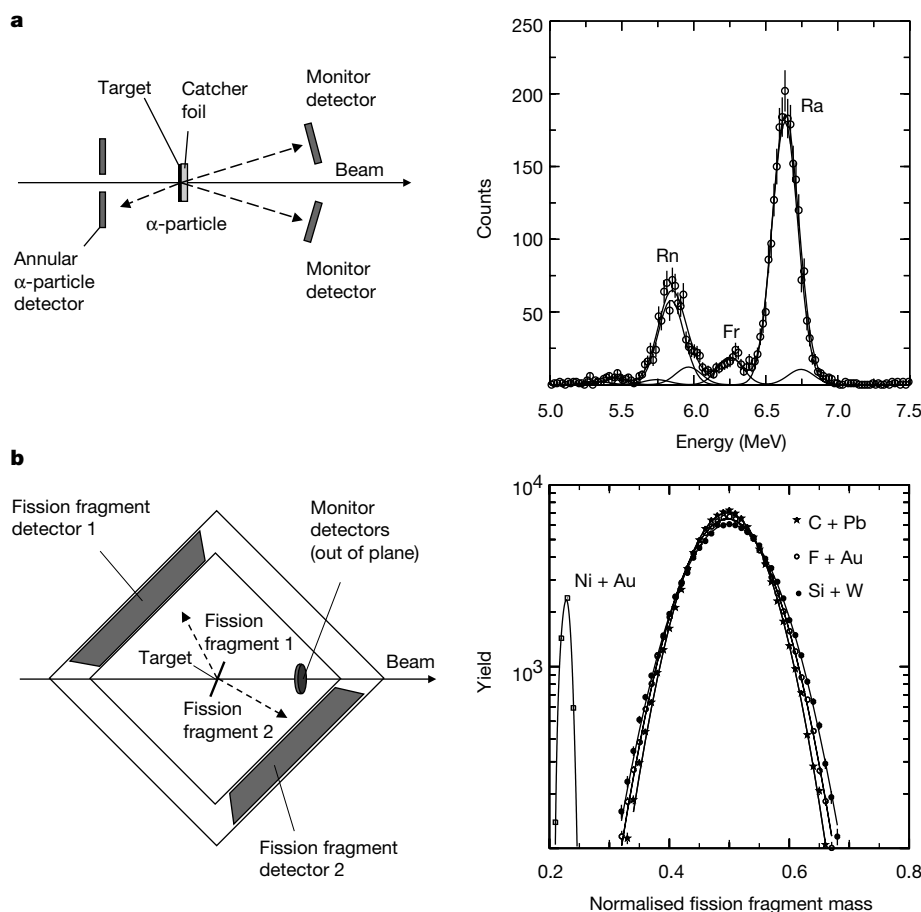
Unstable heavy atomic nuclei not found in nature can be created by fusing two stable nuclei, in a process analogous to colliding charged droplets of liquid. Recently, the formation of a handful of super-heavy nuclei with atomic numbers 114 (ref. 1) and 116 (ref. 2) has been achieved by fusion of heavy nuclei. The electrostatic energy of such systems is very large (which is the reason super-heavy nuclei are unstable), so although the two nuclei may initially be captured by the nuclear potential, rather than fusing, they almost always separate after transfer of mass to the lighter nucleus. This process, called quasi-fission<sup>3,4</sup>, can inhibit fusion by many orders of magnitude. Understanding this inhibition may hold the key to forming more super-heavy elements. Theoretically, inhibition is predicted (ref. 5 and references therein) when the product  $Z_1Z_2$  of the charges of the projectile and target nuclei is larger than about 1,600. Here we report measurements of three fusion reactions with  $Z_1Z_2$  around half this value, each forming

$^{216}_{88}\text{Ra}$ . We find convincing model-independent evidence both of inhibition of fusion, and of the presence of quasi-fission. These results defy interpretation within the standard picture of nuclear fusion and fission.

Several studies have been reported (for example refs 6, 7 and 8) in which compound nuclei were formed by fusing different combinations of projectile and target nuclei. However, our comprehensive and precise measurements, for a carefully chosen range of reactions, are the first to demonstrate the inhibition of fusion for mass-asymmetric reactions.

Three combinations of projectile and target nuclei were studied:  $^{12}\text{C} + ^{204}\text{Pb}$ ,  $^{19}\text{F} + ^{197}\text{Au}$  and  $^{30}\text{Si} + ^{186}\text{W}$ . Following fusion, all form the compound nucleus  $^{216}\text{Ra}$ . We determined the identity and cross-sections of the compound nuclei that survive fission by evaporation of neutrons (and occasionally protons and  $\alpha$ -particles); these nuclei are known as evaporation residues (ERs). We also measured cross-sections for fission, hitherto expected in these reactions to result only from fission following fusion (fusion-fission), but also including contributions from the quasi-fission process (if present) in which no evaporation residues are produced. Measurements were made at ten or more beam energies, from the Coulomb barrier upwards for each reaction. The energies were chosen such that the  $^{216}\text{Ra}$  nuclei were formed at the same excitation energy in the three reactions.

The 16-MV electrostatic accelerator at the Australian National University provided pulsed beams of  $^{12}\text{C}$ ,  $^{19}\text{F}$  and  $^{30}\text{Si}$ , which bombarded thin, isotopically enriched targets. To determine the



**Figure 1** Experimental configurations for the two types of measurement and typical measured spectra. **a**, A schematic diagram of the configuration of the evaporation residue (ER) measurements, with a typical  $\alpha$ -particle energy spectrum from the annular  $\alpha$ -particle detector, and the fit to the spectrum (full lines). This spectrum shows that the  $\alpha$ -particles from Ra can be very clearly identified. **b**, Configuration for the fission measurements,

with three normalized mass spectra (defined as the mass of one fragment divided by the total mass) measured for fission of  $^{216}\text{Ra}$  at the same excitation energy in the three reactions. The peak for elastic scattering of  $^{58}\text{Ni}$  from  $^{197}\text{Au}$  shows the instrumental resolution.

cross-sections for ERs, aluminium catcher foils were placed immediately behind the target to stop the ERs<sup>9</sup>. All the ERs undergo subsequent  $\alpha$ -decay from their ground-states, with lifetimes from 1.9 s to 3.5 h. The decay  $\alpha$ -particles were detected in an annular solid-state detector as shown in Fig. 1a. The catcher foils were 1.5–2.5 times thicker than the average range of the ERs for each reaction. Repeat measurements for the  $^{30}\text{Si}$  reaction with a thicker catcher gave the same cross-sections, confirming that all ERs were stopped.

Measurements of fission were carried out in separate experiments, the two fission fragments being detected in two multi-wire proportional counters with 1-mm position resolution, arranged as shown in Fig. 1b. Each covered an angular range of  $75^\circ$ , allowing determination of the cross-section in one measurement. The coincident detection of both fragments allowed the determination of the fission fragment mass ratio through reconstruction of the velocity ratio<sup>10</sup>. In both ER and fission measurements, relative cross-sections were determined by normalizing to the yields of elastically scattered beam particles, measured in monitor detectors located at  $\pm 22.5^\circ$ . It was vital to obtain reliable absolute cross-sections in order to make meaningful comparisons between reactions. Because the ERs were identified when stopped at the target location, accurate absolute efficiencies could be determined by measuring elastic scattering both in the monitor detectors and in the  $\alpha$ -particle and fission detectors, at low beam energies where the elastic scattering cross-sections follow the Rutherford scattering formula.

The individual ERs, identified by their decay  $\alpha$ -particle energies and lifetimes, were mainly Ra isotopes, formed after the emission of  $x = 3, 4, 5 \dots$  neutrons. The values of  $x$  increase with the initial excitation energy ( $E^*$ ) of the compound nuclei, which in turn is determined by the beam energy and  $Q$ -value. Bohr's independence hypothesis<sup>11</sup> (formulated in 1936 to explain neutron capture) states that compound nuclei with the same excitation energy and angular momentum will decay independently of their method of formation. Our data show that the peak in the ER yield for each isotope occurs at a value of  $E^*$  essentially independent of the reaction, confirming that the actual excitation energies of the  $^{216}\text{Ra}$  nuclei formed in the three reactions are indeed the same.

The sum of the fission and ER cross-sections defines the capture cross-section. In recent years, our understanding of the capture

process in nucleus–nucleus collisions has improved dramatically<sup>12</sup>. The dominant roles played by nuclear deformation and surface vibrations have been shown through the reproduction of precise (1%) capture cross-sections for a wide range of reactions using quantum-mechanical coupled-channels models<sup>13</sup>. Calculations using nuclear potential parameters consistent with many other reactions, and coupling parameters from the known properties of the collective modes in each nucleus, reproduced the present capture cross-sections for the three reactions, typically to within 2%, indicating that the three systems behave as expected.

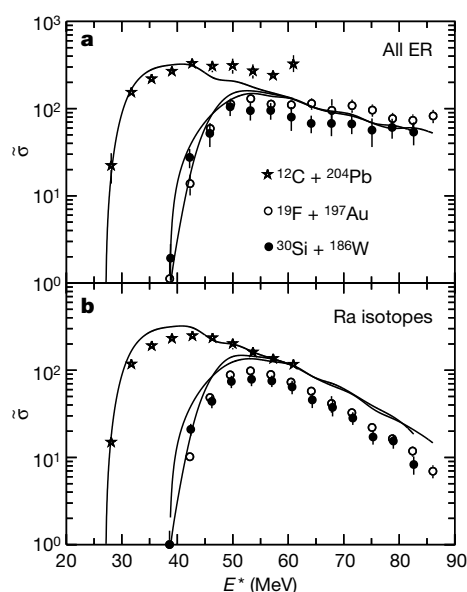
As the beam energy increases, the projectile can bring in higher angular momenta  $l\hbar$ . The height of the fission barrier falls as  $l^2$ , leading to the observed exponential rise in the fission probability. Fitting these data with a statistical decay simulation shows that the fission probability exceeds 98% at  $l = 30$ . This restricts ER survival to low  $l$  values, and allows us to demonstrate below in a simple and essentially model-independent manner that compound-nucleus formation is inhibited for the heavier projectiles.

At a given beam energy  $E$ , the total ER cross-section is the sum of the cross-sections from each  $l$ , these being the products of the capture cross-sections and the probabilities of surviving fission ( $1 - P_f(l, E^*)$ ), thus

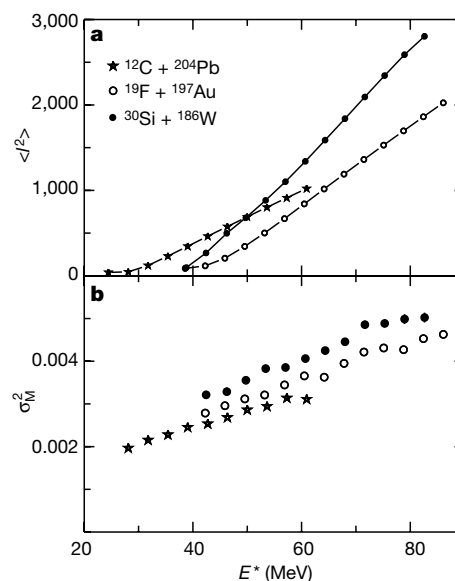
$$\sigma_{\text{ER}}(E) = \pi \lambda^2 \sum_{l=0}^{\infty} (2l+1) \cdot T_l(E) \cdot (1 - P_f(l, E^*)) \quad (1)$$

where  $T_l(E)$  is the probability of capture for  $l$ , and  $\lambda = \lambda/2\pi$  is the reduced de Broglie wavelength, determined by the beam momentum. Division by  $\pi \lambda^2$  gives the reduced cross-section  $\tilde{\sigma}_{\text{ER}}(E)$ . Although the coupled-channels capture models can reliably predict  $T_l(E)$ , reliance on a particular calculation can be eliminated. This is because all such models show that at beam energies sufficiently high above the Coulomb barrier,  $T_l \approx 1$  for all  $l$  which lead to ERs. At these energies  $T_l(E)$  can be dropped, giving

$$\tilde{\sigma}_{\text{ER}}(E) = \sum_{l=0}^{\infty} (2l+1) \cdot (1 - P_f(l, E^*)) \quad (2)$$



**Figure 2** Reduced cross-sections for ERs as a function of excitation energy for (a) all ERs and (b) radium ERs. Monte Carlo statistical model calculations, based on Bohr's independence hypothesis, are shown by the curves, which converge at the higher energies. In contrast, the measured ER-reduced cross-sections for the heavier projectiles are inhibited compared with those for  $^{12}\text{C}$ .



**Figure 3** Experimental evidence for quasi-fission. The calculated mean-squared angular momentum introduced in the three reactions (a), and the measured variance of the normalized fission mass-distributions (b) are shown as a function of excitation energy. Experimental uncertainties are typically smaller than the size of the points. For the same excitation energy and mean-squared angular momentum in  $^{216}\text{Ra}$ , the mass distributions are not the same. The systematic dependence of the variance on the mass asymmetry of the projectile and target is evidence for an increasing contribution from quasi-fission associated with the projectile mass.



Thus at the same  $E^*$ , the three reactions should give the same  $\sigma_{\text{ER}}(E)$ , as long as  $P_l(l, E^*)$  is independent of the reaction, in accordance with Bohr's hypothesis. Monte Carlo statistical model calculations<sup>14</sup> (using  $T_l(E)$  from the coupled-channels calculations that reproduced the measured capture cross-sections) indeed show this behaviour, as illustrated in Fig. 2a, where the three lines converge when all reactions fully populate the angular momenta leading to ERs ( $T_l(E) \cong 1$  for  $l$  below 30). Significantly, the experimental ER data do not converge, but show a systematic difference for each reaction, with the yield of ERs being largest for the lightest projectile. The yields could be affected by the recognized phenomenon of breakup of the lighter projectiles into  $\alpha$ -particle clusters<sup>15</sup>, leading at high  $l$  to incomplete fusion, resulting in increased ER formation rather than fission. Such contributions from incomplete fusion can be rejected by summing the yields for Ra isotopes alone, which can be formed only when all the protons in the projectile are captured. These yields, plotted in Fig. 2b, still show a strong dependence on projectile mass, those for  $^{19}\text{F}$  and  $^{30}\text{Si}$  being respectively  $0.64 \pm 0.09$  and  $0.57 \pm 0.08$  of those for  $^{12}\text{C}$  at the two highest energies. This cannot be explained by low values of  $T_l(E)$ , as the measured capture cross-sections for all three reactions can only be reproduced with  $T_l(E)$  values close to unity. The only reasonable conclusion is that  $P_l(l, E^*)$  depends on the reaction, causing significant inhibition of fusion for the heavier projectiles.

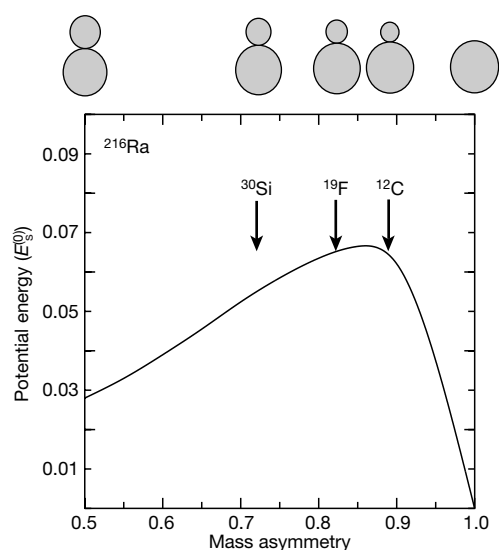
Fusion with light projectiles (such as  $^{19}\text{F}$ ) has previously implicitly been assumed to proceed without any inhibition. We believe that this is the first observation of fusion inhibition for such a light projectile. If the inhibition is caused by the presence of quasi-fission, as in reactions forming super-heavy nuclei, there should be evidence for quasi-fission in the fission mass distributions, because quasi-fission has a broader mass distribution than fusion-fission. The width of the fusion-fission mass distribution depends principally on excitation energy, but may depend slightly on angular momentum. However, if the  $^{216}\text{Ra}$  compound nuclei formed at the same  $E^*$  in different reactions have the same angular momentum distributions, then according to Bohr's hypothesis the fission mass distributions should be identical. The angular momentum distributions can be reliably predicted<sup>16,17</sup> by the coupled-channels calculations that

reproduce the capture cross-sections<sup>12</sup>. We show in Fig. 3a the dependence of the deduced  $\langle l^2 \rangle$  on  $E^*$ , which demonstrates that the  $^{12}\text{C}$  reaction matches the  $^{30}\text{Si}$  reaction at one energy, and approaches the  $^{19}\text{F}$  reaction at the highest energy. Qualitatively this must happen, as the heavier projectiles have a higher threshold  $E^*$  for capture, but carry more angular momentum than  $^{12}\text{C}$ . The measured variances of the fission mass distributions for the three systems, presented in Fig. 3b, do not match at any energy, but rather show a systematic dependence on projectile mass, as well as the expected increase with  $E^*$ . This result is consistent with the observed inhibition of fusion for the heavier projectiles, indicating an increasing contribution from the quasi-fission process, in which no ERs are produced, and for which full mass equilibration is not expected, leading to wider mass distributions for the reactions with heavier projectiles.

These results may be understood qualitatively if we ask what can happen to the projectile nucleus when it makes contact with the target nucleus. It may be swallowed up by the heavier nucleus, resulting in a compact compound nucleus (fusion). Alternatively, it may gain mass from the heavier nucleus, and the system could approach an unstable shape leading to quasi-fission. Models describing certain aspects of these processes have been developed, calculating trajectories over three-dimensional potential-energy surfaces. However, none has predicted fusion inhibition for asymmetric reactions like  $^{19}\text{F} + ^{197}\text{Au}$ . Empirical evidence has led to the suggestion<sup>10</sup> that static deformation of the heavy nucleus affects the competition between fusion and quasi-fission, with an elongated configuration at contact (found only for deformed nuclei) favouring quasi-fission.  $^{204}\text{Pb}$  is essentially spherical, and the small oblate deformation of  $^{197}\text{Au}$  results in contact shapes 2% longer than average at most. The larger prolate deformation of  $^{186}\text{W}$  can result in shapes over 7% longer than average. The largest change in the effect of deformation is thus between  $^{197}\text{Au}$  and  $^{186}\text{W}$ , whereas the largest change in fusion yields is between  $^{197}\text{Au}$  and  $^{204}\text{Pb}$ . Hence this phenomenon seems unlikely to be the main reason for the observed suppression.

The major difference between the three reactions is in their initial mass asymmetry. Fusion and quasi-fission competition has previously been discussed in terms of the dependence of the calculated heights of the mass-asymmetric fission barriers<sup>4,18,19</sup>. These have qualitative similarities to the potential-energy landscape determining the evolution of the system after capture, for which detailed calculations are not available. In Fig. 4 we show the fission barrier energies (interpolated from calculations of ref. 20) for  $^{216}\text{Ra}$  as a function of the mass asymmetry. The projectile will tend to be absorbed if the mass asymmetry is to the right of the peak, whereas if the mass asymmetry is to the left, it will preferentially gain mass. The three reactions studied span the peak, with only the  $^{12}\text{C}$  reaction on the side favouring rapid compound-nucleus formation. It may seem surprising initially that the calculation for the fission shapes agree so well with the experimental data, as the potential maximum for the contact shapes (which are less elongated than the fission shapes) occurs at smaller asymmetries. However, after contact, the development of a neck between the two nuclei causes the potential maximum to shift back to larger asymmetry<sup>21</sup>, similar to that for fission shapes. Our results are consistent with the mass asymmetry being a crucial variable determining the dynamics of these reactions, through the potential energy and the dissipation and inertia tensors. A quantitative understanding of the effect of mass asymmetry and other important variables, such as the deformation of the symmetric fission barrier and the elongation of the system at contact<sup>10</sup>, will require further experimental and theoretical studies.

To describe theoretically the complex multi-dimensional fusion dynamics after capture, diffusion models<sup>22</sup> are being developed. Their future application to fusion reactions leading to new super-heavy elements promises insights into the advantages of using



**Figure 4** Fission barrier energies for  $^{216}\text{Ra}$  shown as a function of mass asymmetry, defined as the difference between the masses of the two parts of the elongated nuclear system divided by their sum. The barrier energies, calculated with the liquid drop model, are expressed in units of the spherical surface energy  $E_s^{(0)}$  of  $^{216}\text{Ra}$ . The initial mass asymmetries for the three reactions are indicated. The shapes at contact corresponding to these asymmetries, and others, are sketched above.

short-lived neutron-rich nuclei with exotic properties, for which extensive experimental surveys will not be practicable. The present comprehensive data set provides a quantitative test of the physics incorporated in these models, by focusing on predictions of the boundary between the simple fusion process envisaged by Bohr, and the complex, dynamical process of quasi-fission. □

Received 28 March; accepted 10 July 2001.

- Oganessian, Yu. Ts. *et al.* Synthesis of nuclei of the superheavy element 114 in reactions induced by  $^{48}\text{Ca}$ . *Nature* **400**, 242–245 (1999).
- Oganessian, Yu. Ts. *et al.* Observation of the decay of  $^{292}116$ . *Phys. Rev. C* **63**, 011301R-1–011301R-2 (2001).
- Back, B. B. Complete fusion and quasifission in reactions between heavy ions. *Phys. Rev. C* **31**, 2104–2112 (1985).
- Toke, J. *et al.* Quasi-fission—the mass-drift mode in heavy-ion reactions. *Nucl. Phys. A* **440**, 327–365 (1985).
- Blocki, J. P., Feldmeier, H. & Swiatecki, W. J. Dynamical hindrance to compound-nucleus formation in heavy-ion reactions. *Nucl. Phys. A* **459**, 145–172 (1986).
- Hinde, D. J., Leigh, J. R., Newton, J. O., Galster, W. & Sie, S. Fission and evaporation competition in  $^{200}\text{Pb}$ . *Nucl. Phys. A* **385**, 109–132 (1982).
- Hahn, R. L., Toth, K. S., LeBeyec, Y., Lagarde, B. & Guidry, M. W. Reactions with  $^{40}\text{Ar}$  and  $^{84}\text{Kr}$  leading to the same compound nucleus,  $^{200}\text{Po}$ . *Phys. Rev. C* **36**, 2132–2135 (1987).
- Sahm, C.-C. *et al.* Fusion probability of symmetric heavy, nuclear systems determined from evaporation-residue cross sections. *Nucl. Phys. A* **441**, 316–343 (1985).
- Morton, C. R. *et al.* Resolution of anomalous fission fragment anisotropies for the  $^{16}\text{O} + ^{208}\text{Pb}$  reaction. *Phys. Rev. C* **52**, 243–251 (1995).
- Hinde, D. J. *et al.* Conclusive evidence for the influence of nuclear orientation on quasifission. *Phys. Rev. C* **53**, 1290–1300 (1996).
- Bohr, N. Neutron capture and nuclear constitution. *Nature* **137**, 344–348 (1936).
- Dasgupta, M., Hinde, D. J., Rowley, N. & Stefanini, A. M. Measuring barriers to fusion. *Annu. Rev. Nucl. Part. Sci.* **48**, 401–462 (1998).
- Hagino, K., Rowley, N. & Kruppa, A. T. A program for coupled-channels calculations with all order couplings for heavy-ion fusion reactions. *Comput. Phys. Commun.* **123**, 143–150 (1999).
- Lestone, J. P. *et al.* Pre-scission charged-particle multiplicities following the reactions  $^{164,167,170}\text{Er} + ^{28}\text{Si}$ . *Nucl. Phys. A* **559**, 277–316 (1993).
- Vergani, P. *et al.* Complete and incomplete fusion and emission of preequilibrium nucleons in the interaction of  $^{12}\text{C}$  with  $^{197}\text{Au}$  below 10 MeV/nucleon. *Phys. Rev. C* **48**, 1815–1827 (1993).
- Rowley, N., Leigh, J. R., Wei, J. X. & Lindsay, R. Obtaining average angular momenta from fusion excitation functions near the Coulomb barrier. *Phys. Lett. B* **314**, 179–184 (1993).
- Balantekin, A. B., DeWeerd, A. J. & Kuyucak, S. Relations between fusion cross sections and average angular momenta. *Phys. Rev. C* **54**, 1853–1862 (1996).
- Ramamurthy, V. S. *et al.* Entrance-channel dependence of fission-fragment anisotropies: a direct experimental signature of fission before equilibration. *Phys. Rev. Lett.* **65**, 25–28 (1990).
- Ravi Prasad, G. V. & Ramamurthy, V. S. Light fragment emission during mass asymmetry relaxation in heavy-ion induced fission. *Phys. Rev. C* **54**, 815–821 (1996).
- Davies, K. T. R. & Sierk, A. J. Conditional saddle-point configurations. *Phys. Rev. C* **31**, 915–922 (1985).
- Möller, P. & Nix, J. R. Potential-energy surfaces for asymmetric heavy-ion reactions. *Nucl. Phys. A* **281**, 354–372 (1977).
- Abe, Y., Ayik, S., Reinhard, P.-G. & Suraud, E. On stochastic approaches of nuclear dynamics. *Phys. Rep.* **275**, 49–196 (1996).

## Acknowledgements

M.D. acknowledges the support of a QEII Fellowship.

Correspondence and requests for materials should be addressed to D.J.H. (e-mail: David.Hinde@anu.edu.au).

# Diffusion of point defects in two-dimensional colloidal crystals

Alexandros Pertsinidis & X. S. Ling

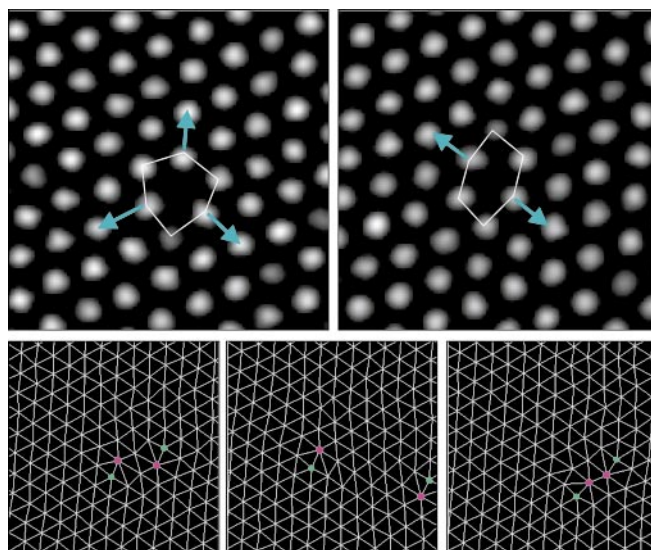
Department of Physics, Brown University, Providence, Rhode Island 02912, USA

Uniform colloidal microspheres dispersed in a solvent will, under appropriate conditions, self-assemble into ordered crystalline structures<sup>1</sup>. Using these colloidal crystals as a model system, a great variety of problems of interest to materials science, physical chemistry, and condensed-matter physics have been investigated during the past two decades. Recently, it has been demonstrated<sup>2</sup> that point defects can be created in two-dimensional colloidal

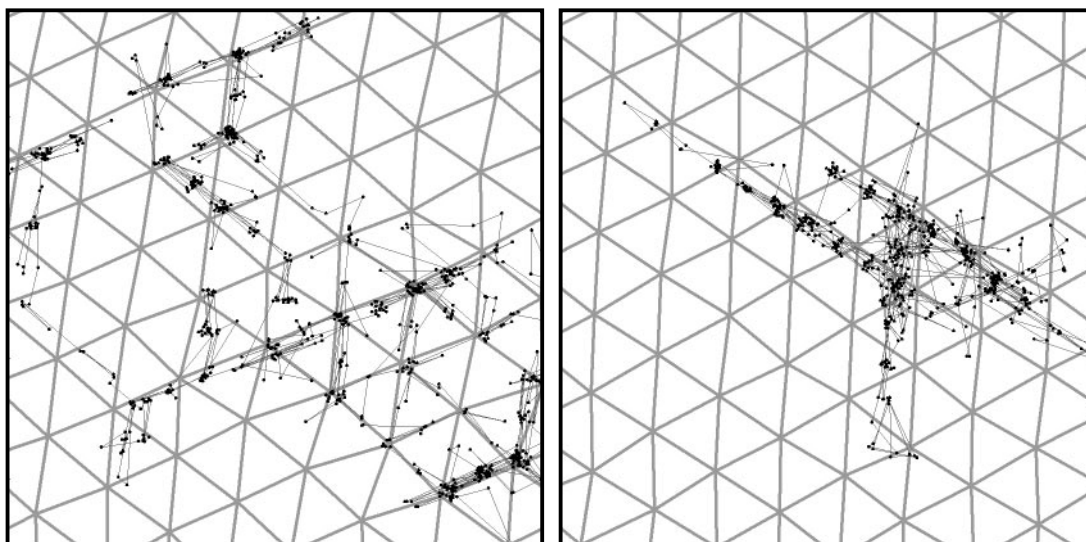
crystals<sup>3</sup> by manipulating individual particles with optical tweezers. Direct imaging of these defects verified that their stable configurations have lower symmetry than the underlying triangular lattice, as predicted by numerical simulations for a number of two-dimensional systems<sup>4–7</sup>. It was also observed that point defects can dissociate into pairs of well-separated dislocations, a topological excitation especially important in two dimensions. Here we use a similar experimental system to study the dynamics of mono- and di-vacancies in two-dimensional colloidal crystals. We see evidence that the excitation of point defects into dislocation pairs enhances the diffusion of di-vacancies. Moreover, the hopping of the defects does not follow a pure random walk, but exhibits surprising memory effects. We expect the results presented in this work to be relevant for explaining the dynamics of other two-dimensional systems.

Our colloidal crystals were made from negatively charged polystyrene-sulphate microspheres (no. 5036, Duke Scientific), 0.360  $\mu\text{m}$  in diameter, suspended in highly deionized water. The particles crystallize owing to strong electrostatic interactions. The experimental details can be found elsewhere<sup>2</sup>. Briefly, a single layer of particles is confined between two fused-silica substrates separated by  $\sim 2 \mu\text{m}$ , resulting in a single-layer colloidal crystal with a lattice constant  $a \approx 1.1 \mu\text{m}$ . The system is quasi-two-dimensional owing to the suppression of the out-of-plane motion of the particles by the negative charge on the silica–water interface. For the conditions of our experiment, the r.m.s. amplitude of the out-of-plane motion is  $87 \pm 5 \text{ nm}$ , measured using a Burleigh Instruments TSE-150 stage (A.P., unpublished results). As a result, the particles do not move out of focus (oil-immersion objective lens, Zeiss Plan-Neofluar, magnification 100 $\times$ , numerical aperture 1.3, depth of focus  $\sim 200 \text{ nm}$ ).

Trapping a particle with optical tweezers<sup>8</sup> and dragging it away from its lattice site creates point defects. We recorded on videotape the dynamics of such single, isolated point defects, using a CCD (charge-coupled device) video camera. Each individual frame (1/60 s) was acquired on a PC and processed using a particle-tracking algorithm<sup>9</sup>. We used a triangulation algorithm to identify the mis-



**Figure 1** Optical micrographs of point defects. Top row, configurations of a mono-vacancy (from ref. 2): left, three-fold symmetric  $V_3$ ; right, crushed  $V_{2a}$ . The 7-fold-coordinated particles are more likely to hop towards the vacant site, resulting in particular easy directions for diffusion (indicated with arrows). Bottom row, di-vacancy breaking into a dislocation pair. Filled circles denote particles that are 5-fold coordinated (green) and 7-fold coordinated (purple). Images are separated by roughly 1/6 s. See text for details.



**Figure 2** Trajectories of defects: left, mono-vacancy; right, di-vacancy. Neighbouring clumps of dots are separated by roughly half the lattice spacing, and correspond to the defect being in a configuration near a lattice point or in a split configuration between two lattice points. The grey lines show the position of the lattice, superimposed on the

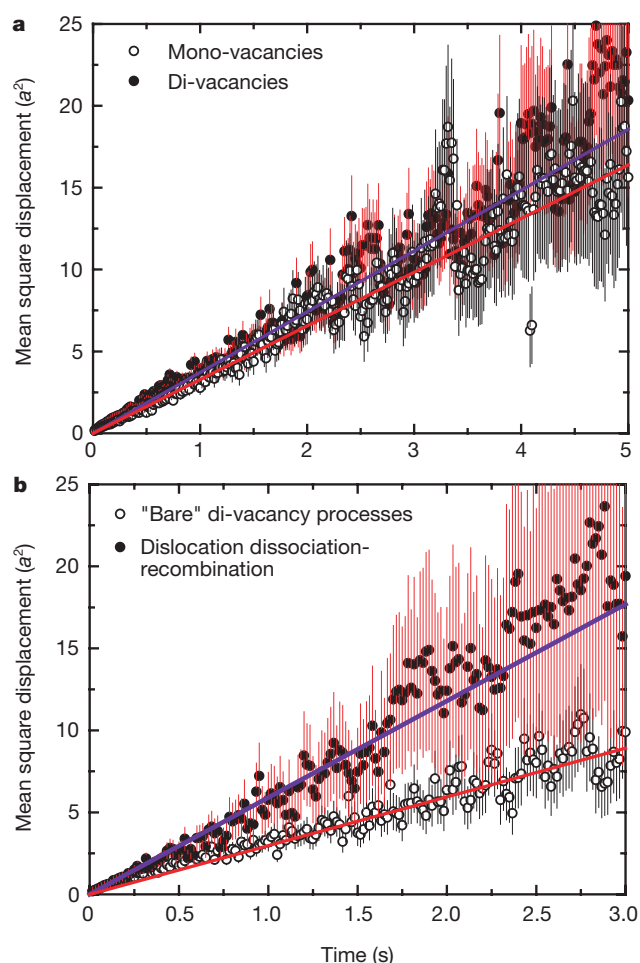
coordinated particles and characterize the configuration of the defect. Figure 1 shows examples of microscopic images of the defects as well as the topological analysis of their configurations. (See also Supplementary Information.)

We tracked a number of mono- and di-vacancies for long enough times (up to 40 s) to obtain reliable measurements of their diffusion constants. Typical trajectories are shown in Fig. 2. The position of the defect  $\mathbf{x}(t)$  is defined as the average of the positions of the 5-fold-coordinated particles around the core of the defect.  $\mathbf{x}(t)$  is determined with a resolution of  $\sim 0.1a$ . Thus we can identify the defect in a particular lattice site or in between two neighbouring sites. To measure the diffusion constants, we take the average of the squared displacements  $\langle \Delta \mathbf{x}^2(\delta t) \rangle$  as a function of the time separation  $\delta t$ , running through all the available trajectories. A linear relationship between  $\langle \Delta \mathbf{x}^2 \rangle$  and  $\delta t$  is obtained for the range 1–5 s (Fig. 3a). The data in this range are weighted by their errors and fitted to a linear relation with a zero  $y$ -intercept, whose slope gives the respective diffusion constants,  $D$ . The results for mono- and di-vacancies respectively are  $D_{\text{mono}}/a^2 = 3.27 \pm 0.03$  Hz and  $D_{\text{di}}/a^2 = 3.71 \pm 0.03$  Hz.

The diffusion constants measured above give the combined contribution of all possible diffusion pathways. In a previous study<sup>2</sup> it was found that the possible stable configurations of mono- and di-vacancies are not just the ‘well-identified’ point defect configurations, but that there are also configurations where the point defects appear as separated dislocation pairs. Here we present evidence that these topological excitations are important in the diffusion of point defects. We observe that the defects dissociate into two dislocations that glide from site to site and finally recombine at a new position that can be several lattice constants away from the original one. The individual dislocations glide faster than the hopping of mono- or di-vacancies. Such a process is a distinct diffusion pathway, giving a separate contribution ( $D_{\text{pair}}$ ) to the diffusion constant.

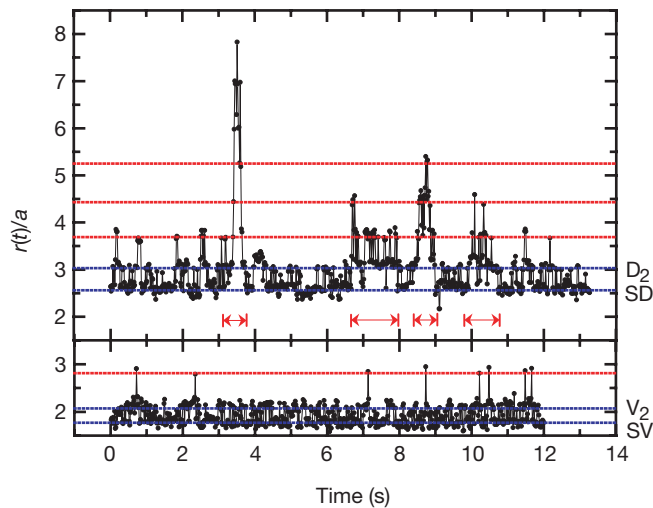
We found that mono-vacancies spend considerably less time as a dislocation pair than do di-vacancies, and the above diffusion mechanism is not important for mono-vacancies. The different behaviour exhibited by mono- and di-vacancies can be easily distinguished, as illustrated in Fig. 4, where the separation of the dislocations comprising the defect is shown as a function of time. The mono-vacancy is observed primarily to fluctuate between

trajectories a few seconds after the defect diffused away from the region shown. The positions of the defects are displaced with respect to the lattice, owing to slow, long-wavelength, thermally generated deformations.



**Figure 3** Time dependence of mean square displacements. **a**, Mono- and di-vacancies (open circles with black error bars and filled circles with red error bars, respectively). The average of  $\langle \Delta \mathbf{x}^2 \rangle$  as function of  $\delta t$  was measured from  $\sim 2$ -minute trajectories. The red and purple straight lines are linear fits for mono- and di-vacancies, respectively, for  $\delta t$  in the range of 1–5 s. **b**,  $\langle \Delta \mathbf{x}^2 \rangle$  separately for ‘bare’ di-vacancies and for di-vacancies excited into dislocation pairs. Fits are for  $\delta t$  values of 1–3 s.



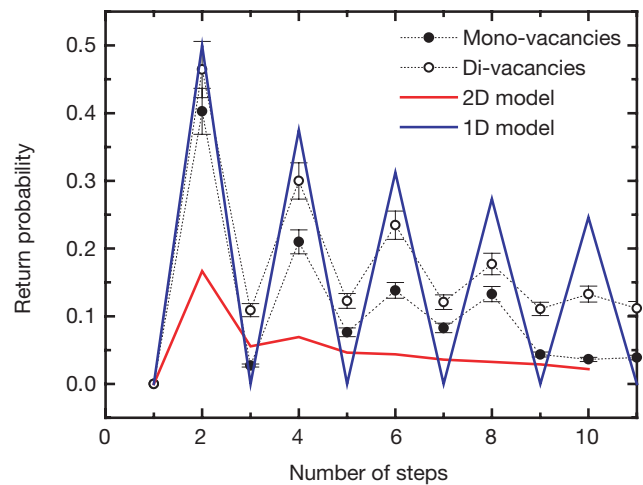


**Figure 4** Separation  $r(t)$  of the dislocations comprising a mono-vacancy (top trace) and a di-vacancy (bottom trace).  $r(t)$  is measured as the standard deviation of positions of the 5-fold-coordinated particles in the defect core. Dotted horizontal lines indicate the positions of the minima in the dislocation pair potential, as found in ref. 2. The blue lines correspond to configurations of bound dislocations. These are the split and crushed configurations (labelled SV, SD and  $V_2$ ,  $D_2$ , respectively) of mono- and di-vacancies, discussed in detail in ref. 2. The red lines designate configurations of two well-separated dislocations. In the case of di-vacancy, the arrows indicate intervals in which the defect appears as a repeatedly dissociating and recombining dislocation pair.

configurations in which the dislocations are bound, with very rare excursions into configurations where the dislocations dissociate. In contrast, the di-vacancy frequently spends long intervals (a few seconds) as a dislocation pair, with the two dislocations dissociating and recombining several times during this period.

The combined diffusion constant for di-vacancies is  $D_{\text{di}} = pD_{\text{pair}} + (1-p)D_{\text{di}}^0$ , where  $p \approx 0.25$  is the fraction of time the di-vacancies are excited into a dislocation pair, and  $D_{\text{pair}}$  and  $D_{\text{di}}^0$  are the respective contributions from the dissociation–recombination mechanism and from the ‘bare’ di-vacancy hopping. In order to estimate the diffusion constants for the two different mechanisms, we link into two separate trajectories the intervals spent as a dislocation pair and as a ‘bare’ di-vacancy. To avoid systematically picking up hopping events, the shortest interval was 1 s (larger than the time between jumps). Carrying out the same procedure as for the original trajectories, we found the two diffusion constants to be  $D_{\text{pair}}/a^2 = 5.90 \pm 0.09$  Hz and  $D_{\text{di}}^0/a^2 = 2.97 \pm 0.03$  Hz (Fig. 3b). These numbers combine to give the measured value  $D_{\text{di}}/a^2 = 3.70$  Hz.

We now turn our attention to the details of the underlying stochastic hopping processes. The trajectories of the defects would be naively expected to be pure random walks on a triangular lattice—that is, when hopping, the system chooses with equal probability between the six neighbour sites and consecutive events are uncorrelated. However, a careful analysis reveals a peculiar feature: unlike a two-dimensional (2D) random walk, the motion of the defects is often restrained along particular directions. The existence of easy directions for diffusion is related to the low symmetry of the stable configurations of the relaxed lattice around the defect. The probability for hopping towards the six nearest neighbours is not the same for all of them, but is higher along particular lines of symmetry (Fig. 4). For di-vacancies excited into a pair of separated dislocations, the easy direction coincides with the gliding direction of the two dislocations. The easy direction for diffusion remains unchanged for up to a few seconds, that is, several consecutive hopping events.



**Figure 5**  $P_0(n)$  measured from 40 s of data, or roughly 60 hopping events. The behaviour observed for both type of defects is more like that predicted for a 1D, rather than a 2D, random walk. The data show oscillations, with a period of two steps, and the average trend lies between the curves for 1D and 2D walks.

The existence of easy directions for hopping makes the trajectories of the defects quasi-one-dimensional (quasi-1D). We can demonstrate this effect by measuring the probability  $P_0(n)$  for a defect to return to its original position after it has executed  $n$  steps. We found that the measured  $P_0(n)$  for small  $n$  resembles the value for random walk on a 1D chain, rather than that for random walk on a 2D triangular lattice. This is illustrated in Fig. 5:  $P_0(n)$  clearly deviates from the expected 2D behaviour, following the characteristic oscillations of the 1D random walk with a period of two steps.

Another surprising property of the diffusion of these defects involves memory effects during hopping. The defects appear to spend some time around a lattice site, attempting to hop towards neighbouring sites by changing into a split intermediate configuration between the two lattice points. Depending on the symmetry of the relaxed lattice in the original configuration (2- or 3-fold), the defect is seen to attempt hopping towards either two or three of the six neighbours. The defect typically hops back to the original site, finally getting to the next site only after several attempts. The back-hopping to forward-hopping ratio is found to be  $\sim 3:1$ .

We suggest that this memory effect is a result of the slow relaxation rates of distortions in the over-damped colloidal crystal. The core region of the defect can rearrange rapidly ( $\sim 10$  Hz), but the rest of the lattice around the defect cannot follow at the same rate (for example, distortions with wavelength  $10a$  relax at  $\sim 1$  Hz). As a vacancy changes from a configuration at one site to a split intermediate configuration between two sites, the rest of the lattice does not immediately relax, and seems to ‘push’ the vacancy back to its initial site.

The results we present here are relevant to a number of other systems, such as Wigner electron crystals in low-carrier-density semiconductor heterostructures and Abrikosov vortex lattices, formed by lines of magnetic flux quanta in type-II superconductors. In particular, dynamical processes in the vortex lattice should resemble those we observed in the colloidal crystal—as the dynamics of vortices in type II superconductors are also over-damped. Our observations might also apply to the dynamics of surfaces of atomic crystals. □

Received 28 March; accepted 21 June 2001.

- Gast, A. P. & Russel, W. B. Simple ordering in complex fluids. *Phys. Today* **51**(12), 24–30 (1998).
- Pertsinidis, A. & Ling, X. S. Equilibrium configurations and energetics of point defects in two-dimensional colloidal crystals. *Phys. Rev. Lett.* **87**, 098303 (2001).

3. Murray, C. A. in *Bond-Orientational Order in Condensed Matter Systems* (ed. Strandburg, K. J.) Ch. 4 (Springer, New York, 1992).
4. Fisher, D. S., Halperin, B. I. & Morf, R. Defects in the two-dimensional electron solid and implications for melting. *Phys. Rev. B* **20**, 4692–4712 (1979).
5. Cockayne, E. & Elser, V. Energetics of point defects in the two-dimensional Wigner crystal. *Phys. Rev. B* **43**, 623–629 (1991).
6. Frey, E., Nelson, D. R. & Fisher, D. S. Interstitials, vacancies, and supersolid order in vortex crystals. *Phys. Rev. B* **49**, 9723–9745 (1994).
7. Jain, S. & Nelson, D. R. Statistical mechanics of vacancy and interstitial strings in hexagonal columnar crystals. *Phys. Rev. E* **61**, 1599–1615 (2000).
8. Ashkin, A., Dziedzic, J. M., Bjorkholm, J. E. & Chu, S. Observation of a single-beam gradient force optical trap for dielectric particles. *Opt. Lett.* **11**, 288–290 (1986).
9. Crocker, J. C. & Grier, D. G. Methods of digital video microscopy for colloidal studies. *J. Colloid Interf. Sci.* **179**, 298–310 (1996).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

## Acknowledgements

We thank S.C. Ying and D.A. Weitz for discussions. This work was supported by the National Science Foundation, the Petroleum Research Fund, and the Research Corporation. X.S.L. acknowledges the support of the A.P. Sloan Foundation.

Correspondence and requests for materials should be addressed to A.P. (e-mail: [pertsin@barus.physics.brown.edu](mailto:pertsin@barus.physics.brown.edu)).

## Evolution of magma-poor continental margins from rifting to seafloor spreading

R. B. Whitmarsh\*, G. Manatschal† & T. A. Minshull\*

\* Southampton Oceanography Centre, European Way, Southampton, SO14 3ZH, UK

† CGS-EOST, Université Louis Pasteur, 1 rue Blessig, 67084 Strasbourg, France

The rifting of continents involves faulting (tectonism) and magmatism, which reflect the strain-rate and temperature dependent processes of solid-state deformation and decompression melting within the Earth<sup>1,2</sup>. Most models of this rifting have treated tectonism and magmatism separately, and few numerical simulations have attempted to include continental break-up and melting, let alone describe how continental rifting evolves into seafloor spreading. Models of this evolution conventionally juxtapose continental and oceanic crust. Here we present observations that support the existence of a zone of exhumed continental mantle, several tens of kilometres wide, between oceanic and continental crust on continental margins where magma-poor rifting has taken place. We present geophysical and geological observations from the west Iberia margin<sup>3–7</sup>, and geological mapping of margins of the former Tethys ocean now exposed in the Alps<sup>8–13</sup>. We use these complementary findings to propose a conceptual model that focuses on the final stage of continental extension and break-up, and the creation of a zone of exhumed continental mantle that evolves oceanward into seafloor spreading. We conclude that the evolving stress and thermal fields are constrained by a rising and narrowing ridge of asthenospheric mantle, and that magmatism and rates of extension systematically increase oceanward.

The west Iberia–Newfoundland conjugate margins experienced a final Early Cretaceous phase of rifting and seafloor spreading that started ~133 Myr ago<sup>7</sup>. The Alpine Tethyan margins experienced a final Late Triassic/Early Jurassic rift phase presaging the Liguria–Piemonte ocean where seafloor-spreading magmatism began 160–165 Myr ago<sup>13</sup>. Both the west Iberia and Alpine magma-poor

margins are characterized by margin-parallel deep-water zones, apparently representing successive stages of margin evolution, that is, thinned continental crust dissected by low-angle detachment faults succeeded by exhumed sub-continental mantle, with oceanward-increasing mafic melt volumes, that merges into oceanic crust<sup>3,6,13,14</sup>. We describe each zone in turn, adducing evidence from margins in both areas, before presenting models to explain the observations.

The west Iberia margin exhibits tilted fault blocks bounded by west-facing normal faults<sup>5,14,15</sup>. Under the slope and rise larger blocks lie 10–20 km apart over continental crust 20–30 km thick. Oceanward, wherever crustal blocks are ≤7 km thick, such faults are clearly listric (downward-flattening and concave upward), even penetrating the mantle, and closer together; the crust tapers oceanward to zero thickness within a distance of 50 km (Figs 1d and 2)<sup>3–5,15,16</sup>. Such blocks, where drilled by the Ocean Drilling Program (ODP), are capped by late Tithonian (~146 Myr ago) shallow-water sediments and therefore are continental crust<sup>7</sup>. Thus a marked change occurs in the mechanical response of the continental lithosphere to extension once the crust has been thinned to less than about 7 km.

In the southern Iberia abyssal plain (SIAP) three adjacent ODP boreholes (Sites 900, 1067 and 1068; Figs 1c and 2), near where the crust tapers to zero, yielded gabbro, amphibolite, tonalite gneiss, anorthosite (all part of a <400-m-thick basement layer floored by a sub-horizontal tectonic crust/mantle contact<sup>16</sup>) and serpentinized peridotite. The basement appears to be capped by a younger low-angle detachment fault that breaks out 6 km to the east and plunges into basement 14 km to the west<sup>10</sup>. The layer contains pre-rift lower crustal rocks, underplated in Late/post-Hercynian time (270 ± 3 (± 1σ) Myr ago), that later underwent ductile deformation (at 0.6–0.8 MPa) and cooling before exhumation at the seabed (<137 Myr ago)<sup>10</sup>.

In the Alps, the continental crust of the distal Adria margin appears as fault blocks and as isolated allochthons (masses of rock that have been tectonically moved from their places of origin) separated by sub-horizontal detachments from underlying mantle<sup>11</sup>. This crust is mainly composed of Late to post-Hercynian granites intrusive into poly-metamorphic basement. Sporadically, underplated meta-gabbros exhumed from a pre-rift lower crustal level are observed which intruded the crust–mantle boundary during Late/post-Hercynian time<sup>17</sup>. Pressure–temperature–time (*P–T–t*) data indicate that the crust–mantle boundary was rather cool (~550 °C, 0.9–1.0 MPa) at the onset of rifting<sup>18</sup>. All structures genetically linked to the final phase of crustal extension, leading to formation of the zone of exhumed continental mantle (ZECM), are brittle structures.

Thus, the distal continental parts of the margins are dissected by listric detachment faults which either separate continental crustal blocks from mantle or occasionally penetrate the mantle. Where pre-rift lower crustal rocks are exhumed they frequently show mafic compositions and similar pre-rift *P–T–t* evolutions beginning with crustal underplating followed by isobaric cooling to 550 °C. Syn-rift melt products (contemporaneous with the rifting process) are absent. We estimate that durations of continental extension were, respectively, a few tens of Myr (at <5 mm yr<sup>–1</sup>) for the Tethyan margins<sup>19</sup> and ~10 Myr (final phase, at >3.5 mm yr<sup>–1</sup>) off Iberia<sup>20</sup>.

The ZECM in the SIAP is 40- to 170-km wide with distinctive geophysical characteristics<sup>3,4</sup>. Morphologically, the basement surface identified on seismic reflection profiles may be divided into two regions of N–S-trending basement ridges and deep (~9 km) relatively low-relief basement (Fig. 1a and c); both narrow northward. Highly serpentinized peridotite was cored at four ODP sites mostly near the margins of the ZECM<sup>7,21,22</sup>. Primary-phase chemistry and clinopyroxene trace-element compositions indicate heterogeneous mantle depleted by heterogeneous, less than 10%,

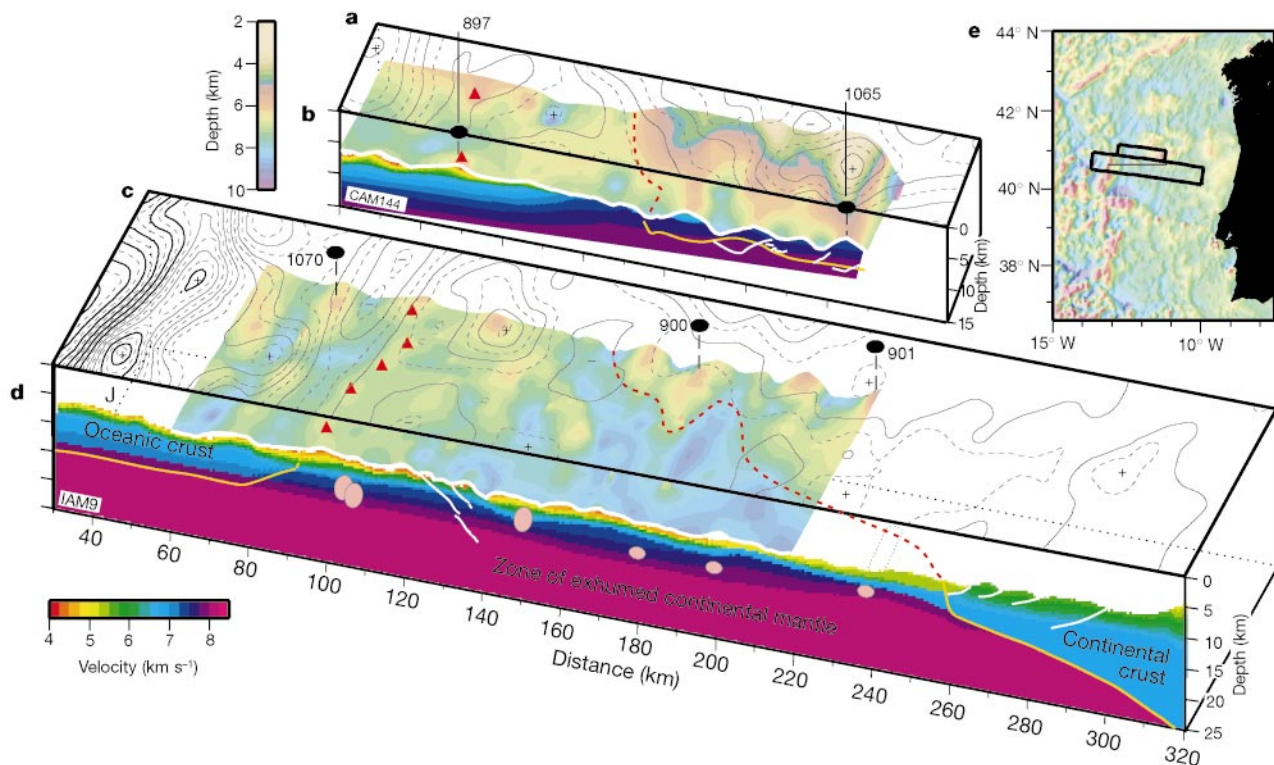
partial melting, and percolated by mafic melts<sup>21</sup>. Trace-element compositions are more like sub-continental or suprasubduction-zone mantle than abyssal (oceanic) mantle<sup>21,22</sup>. Serpentinization began, at least locally, before seabed exhumation of the peridotite<sup>23</sup>.

The SIAP ZECM has a seismic velocity structure which differs from those of the adjacent stretched continental and oceanic crusts (Fig. 1d)<sup>3</sup>. A 2–4-km-thick upper basement layer with a P-wave velocity of 4.5–7.0 km s<sup>-1</sup> and a high velocity gradient (~1 s<sup>-1</sup>) merges into a lower layer not more than 4 km thick with velocities of ~7.6 km s<sup>-1</sup> and a low velocity gradient (<0.2 s<sup>-1</sup>). Moho reflections are weak or absent. The top-basement velocity is lower than that of the adjacent continental crust, and the velocity in the lower layer is too high to represent magmatically intruded or underplated continental crust or even oceanic layer 3 (the lower oceanic crust)<sup>3</sup>. Therefore the velocity structure probably reflects decreasing mantle serpentinization with depth<sup>3–5</sup>. Low-amplitude N–S magnetic anomalies (Fig. 1a, c and e) indicate that ZECM basement magnetizations are typically much lower than those of oceanic basement further west, and the spectral properties of the anomalies suggest that most source bodies, equated here with syn-rift mafic intrusions (for example, ref. 24), lie not at top basement but up to 6.5 km deeper<sup>6</sup>. These observations suggest that the upper seismic layer contains little magmatic material, whereas the lower layer contains isolated N–S elongated magmatic intrusions which increase in volume oceanward (Fig. 1d). Although numerical models predict the addition of 3–5 km of melt to the ZECM we infer that, because

the region of asthenospheric upwelling was relatively wide at this stage, very little decompression melting took place until the mantle viscosity structure approached that inferred to exist under a mid-ocean ridge<sup>20</sup>.

In the Alps the sub-continental affiliation of mantle rocks is indicated by clinopyroxene Sr and Nd isotope compositions similar to many western Mediterranean sub-continental orogenic spinel lherzolites<sup>25</sup>, by association with lower continental crust<sup>26</sup> and by isolated continental allochthons overlying exhumed mantle<sup>12</sup>. Sedimentary infill structures and serpentine clasts in post-rift Upper Jurassic deep-sea sediments indicate submarine exhumation. Deeper mantle rocks are encountered oceanwards. At the Platta nappe (sheet of rock thrust sideways over adjacent strata), gabbros 160.9 ± 0.5 Myr old were intruded into an already serpentinized peridotite<sup>13</sup> and provide direct evidence of syn-tectonic ZECM magmatism. Intrusion was immediately followed by emplacement of massive mid-ocean ridge basalt (MORB) pillows marking the onset of seafloor spreading<sup>13</sup>.

Off Iberia, low-angle intrabasement ZECM faults have not been detected and deeply penetrating landward-dipping high-angle faults are imaged only rarely<sup>5</sup>. In contrast, Alpine geology, supported by top-basement tectono-sedimentary breccias in several ODP boreholes<sup>7</sup>, reveals that the basement surface often represents the footwall of a series of mantle detachment faults individually accommodating ~10–20 km offsets in which the hanging wall moved oceanward. Because the faults cap basement highs that have minor submarine relief, despite their large offset, and enlisting



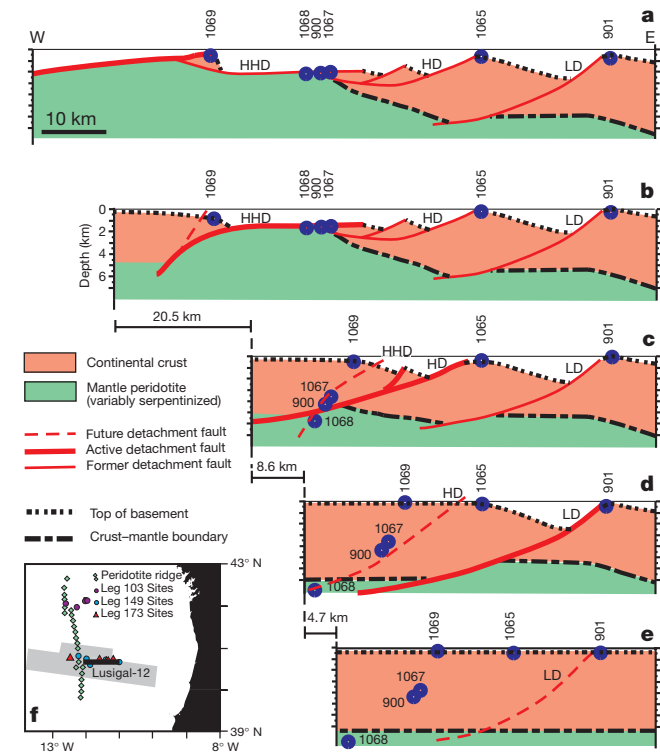
**Figure 1** Exploded block diagram of the continental margin beneath the SIAP off west Iberia. **b, d**, Cross-sections of seismic profiles CAM144 (**b**, ref. 4) and IAM9 (**d**, ref. 3) colour-coded in km s<sup>-1</sup> with the crust–mantle boundary in yellow-brown (where resolvable). **a, c**, Plan views of the three margin-parallel regions of thinned continental crust (bounded by dashed red line), the zone of exhumed continental mantle (ZECM) and the oceanic crust (roughly bounded by peridotite ridge, red triangles)<sup>7</sup>. The plan views also include relief of acoustic basement (depth below sea-level), contoured sea-surface magnetic anomalies (25 nT contours; crosses and dashes denote respectively peaks and troughs)<sup>7,16</sup> and ODP Sites 1070, 897, 900, 1065 and 901 (dots). Detachment faults

(**b**; CAM144 thin white lines<sup>16</sup>) evolved as in Fig. 2, other deep faults in the ZECM (**d**; IAM9 white lines<sup>5</sup>) are suspected to be originally sub-horizontal mantle boundaries (Fig. 3c). The southern block illustrates how gabbroic margin-parallel magnetic source bodies (pink ellipses in **d**)—intruded within the ZECM and inferred to exist from magnetic anomaly inversions<sup>6</sup>—become more prevalent and closely spaced oceanwards, evolving eventually into a continuous layer of intrusions characteristic of ocean layer 3. **e**, Block locations, Lusigal-12 track (Fig. 2) and shaded-relief sea-surface magnetic anomalies.



other arguments (see Fig. 2b legend; ref. 10), we suggest that mantle exhumation was accomplished by concave-downward faults shortly before the onset of seafloor spreading.

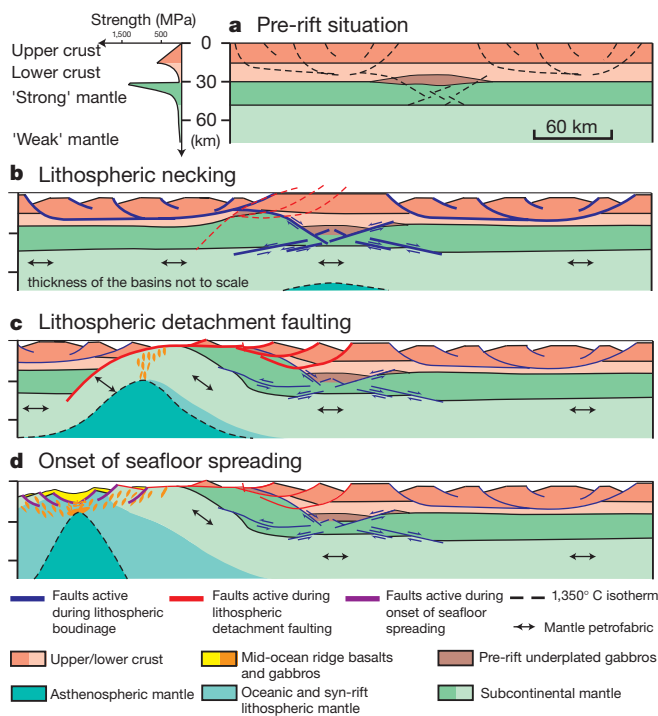
Off west Iberia the ZECM merges into the earliest oceanic crust near a margin-parallel basement peridotite ridge<sup>14</sup> (Fig. 1d). High-amplitude linear margin-parallel magnetic anomalies immediately west of the peridotite ridge in the SIAP are consistent with seafloor spreading at  $10 \text{ mm yr}^{-1}$ , beginning around 126 Myr ago<sup>27</sup>. The seismic velocity structure gradually changes oceanwards to typically oceanic 10–20 km west of the peridotite ridge<sup>3</sup>. At ODP Site 1070, 20 km (representing  $\sim 2 \text{ Myr}$  of spreading) oceanward of



**Figure 2** Palinspastic reconstruction with arbitrary reference datum<sup>10</sup> of profile Lusigal 12 (ref. 16) across the southern Iberia abyssal plain (SIAP). **a**, Tectonic interpretation of the depth-migrated multichannel seismic reflection profile. The total extension accommodated by the detachment faults LD, HD and HHD is around 34 km and this occurred between late Tithonian time and 137 Myr ago (that is, over a period of  $\leq 9 \text{ Myr}$ ) at  $> 3.5 \text{ mm yr}^{-1}$ . Exhumation of about 75 km of subcontinental mantle within the ZECM at not less than  $7.2 \text{ mm yr}^{-1}$  was mainly accommodated by concave-downward faults<sup>10</sup>. **b**, Situation after mantle exhumation along the detachment HHD. The roll-over of HHD can explain three characteristic features: (1) the large (20.5 km) offset on this fault resulting in the observed minor submarine relief within most of the ZECM; (2) exhumation of subcontinental mantle by pulling it out from underneath a relatively stable hanging wall; and (3) the formation of tectono-sedimentary breccias on two basement highs drilled by the ODP (Sites 899 and 900/1067/1068) by a conveyor-belt-type sediment accumulation whereby the exhumed footwall rocks were fractured, exposed, and redeposited along the same active fault system. **c**, Situation at the onset of HHD and after motion on detachment HD stopped. We note the positions of Sites 900 and 1067 in the footwall of the later HHD; identical plagioclase  $^{40}\text{Ar}/^{39}\text{Ar}$  ages (137 Myr ago) at both sites demonstrate that they cooled and were exhumed at the sea floor simultaneously. Thus, the transition from Fig. 2c to b had to occur about 137 Myr ago. **d**, Situation after extension along the detachment LD. We note the listric geometry of this fault, the relatively small 4.7 km offset along it, compared to the later detachment HHD, and its extension into brittle mantle. **e**, Situation at the onset of detachment faulting when continental crust was already thinned over a broad region by symmetric lithospheric-scale pure shear from about 30–35 km to about 7 km, showing distribution of the rocks eventually drilled by ODP. **f**, Track of profile Lusigal 12 and areas of Fig. 1 blocks.

the peridotite ridge (Fig. 1d), we cored Late Aptian (112.2–116.9 Myr ago) sediments over pegmatite gabbro and depleted subcontinental serpentinized peridotite with gabbro veins<sup>21,22</sup> without encountering basaltic rocks. The gabbro had an enriched mid-ocean-ridge basalt (E-MORB) source<sup>21</sup>. The pegmatite cooled below  $500^\circ\text{C}$  at  $119 \pm 0.7 \text{ Myr ago}^{10}$ . Thus, although near Site 1070 we encounter geophysical characteristics of normal oceanic crust, the cores reveal subcontinental mantle that was intruded by MORB dykes, exhumed at the seafloor and then sedimented shortly afterwards. The dates imply that tectonism of the earliest ‘oceanic’ crust continued after its emplacement over a region tens of kilometres wide.

In the Alps, oceanic crust is represented by undeformed MORB pillow lavas that thicken oceanwards. Generally, just oceanward of the edge of the continental crust, basalts form isolated bodies less



**Figure 3** Conceptual lithosphere-scale model of development of a rifted magma-poor margin relative to a fixed right-hand edge. **a**, Initial situation with four-layer rheology (justified because the pre-rift lower crust often contains quartz-rich horizons<sup>7,26</sup>) and crust locally thickened by pre-rift underplated gabbro. **b**, Initially the upper mantle, the strongest part of the lithosphere, necked beneath the gabbros where it was weakest, allowing the asthenosphere to rise. Elsewhere, ductile flow of the lower crust determined where rift basins formed at the surface. Relatively little crustal thinning and subsidence occurred above the necked region, whereas the adjacent areas were strongly thinned. Arrows indicate a hypothetical, initially sub-horizontal, petrofabric. **c**, Later, the thermal structure and gravitational response associated with the rising asthenosphere started to influence the rifting. Then, rifting was localized at the margin of the relatively weak unextended, or only slightly extended, crust. This allows the important initial thinning of the lower crust and its observed abrupt transition to weakly thinned crust (see for example refs 17 and 26) to be explained. The change in extensional geometry from listric to one or more concave-downward faults reflected a change in the distribution of weak layers. Whereas listric faults sole out in horizontal weak layers, concave-downward faults might be favoured by sub-vertical weak zones, possibly resulting from rising magma and higher thermal gradients above the rising and narrowing asthenosphere. **d**, The asthenosphere ascended close to the surface and mid-ocean-ridge basalt melts were intruded into, and even extruded onto, sub-continental mantle. Deeper mantle layers were exhumed oceanward. Eventually, increasing melt production led to the creation of ‘continuous’ oceanic crust and the asthenosphere spawned oceanic upper mantle. Seafloor spreading had begun.

than 100 m thick; further oceanwards, basalt bodies are aligned parallel to the margin. The first voluminous basalts show a clearly MORB isotopic signature<sup>13</sup> and overlie tectonically exhumed sub-continental mantle.

Thus, generally, after the break-up of continental crust, mafic intrusives and/or extrusives increase in volume oceanwards across the ZECM. They have a MORB signature, yet intrude and overlie sub-continental mantle. Eventually, such a 'crust' can acquire a geophysically oceanic signature possibly even before extrusives/intrusives of oceanic layer 2 (upper oceanic crust) are widespread.

We now explain how the upper ~7 km of the ZECM evolves into oceanic crust formed by seafloor spreading, before we present a broader lithospheric-scale perspective. In a late stage of rifting, when the continental crust was already locally thinned to about 7 km (Fig. 2e), extension became focused within the future distal margin and was accommodated by detachment faulting (Fig. 2d to b). Exhumation of sub-continental mantle was mainly accommodated by concave-downward faults (Fig. 2b to a). The crust and uppermost mantle underwent brittle deformation and extension was dominated by simple shear. Seafloor spreading, heralded by an oceanward-increasing emplacement of MORB melt over, and into, depleted sub-continental mantle, eventually generated a thin oceanic crust (Fig. 1d). Tectonism remained active over a broad zone at least until slow seafloor-spreading had started. Extension rates increased by factors of 2–3 from the latest continental rifting to the earliest seafloor spreading.

Our conceptual lithospheric-scale model (Fig. 3) exhibits the well known sequential modes of extension from pure shear to simple shear to slow seafloor-spreading. However, only now can we characterize the simple-shear phase, infer its 'genetic' link to the preceding pure-shear phase and understand its link to the subsequent seafloor-spreading phase. Rifting began with an isostatically equilibrated ~30-km-thick continental crust (Fig. 3a). Initial rifting was controlled by ductile flow of the lower crust, as suggested by some numerical models<sup>28</sup>, and associated with subsidence<sup>9</sup>. Rift basins were bounded by listric faults that levelled out at mid-crustal levels, where the crust was weaker, and decoupled upper crust and upper mantle deformation. Rifting was distributed over the entire future margin (Fig. 3b). The extending crust cooled during ongoing rifting. Magmatism was essentially absent during this initial stage of rifting and the rift was, on a lithospheric scale, symmetric.

Mechanically or thermally induced weaknesses in the upper mantle, possibly resulting from pre-rift underplating, controlled the initial location of upper mantle necking which can be offset from crustal deformation, as in some numerical models (see, for example, ref. 28). The weaknesses guided the ascent of the asthenosphere, which in turn controlled the location and evolution of the latest rifting and eventual continental break-up (Fig. 3c). The transition from symmetric non-magmatic rifting to seafloor spreading included a transient phase of simple-shear dominated, and partly asymmetric, continental rifting that coincided with the onset of systematically increasing magmatism. This transition was characterized by the localization of extension along continental crust and even upper mantle detachment faults and then along top-to-the-ocean and concave-downward faults, leading to the exhumation of the ZECM (Fig. 3c). Finally, the ascending but narrowing asthenospheric mantle (Fig. 3d) led to the typical focused magmatism and tectonism of a mid-ocean ridge.

Although our model involves mechanisms of tectonic deformation that are not yet fully understood or proved, it does explain a wide range of observations. It predicts low-angle detachments in continental crust (Fig. 2a), a systematic oceanward increase in synkinematic MORB extrusives/intrusives across the ZECM (Fig. 3d) and a systematic trend from shallow to deeper exhumed sub-continental mantle rocks across a reconstructed conjugate pair of ZECMs in association with a rotation of mantle rock structures (Fig. 3c).

Other authors have suggested that other magma-poor rifted margins in the North Atlantic<sup>29</sup> and elsewhere have characteristics, including ZECMs, that are at least partially similar to those presented above. Elsewhere, for example, in the Woodlark Basin<sup>30</sup>, the transition from continent to ocean appears to be more abrupt. Further observations may resolve whether these differences result from differing rates and durations of extension and/or other factors. □

Received 19 March; accepted 19 July 2001.

- McKenzie, D. P. & Bickle, M. J. The volume and composition of melt generated by extension of the lithosphere. *J. Petrol.* **29**, 625–679 (1988).
- England, P. Constraints on extension of continental lithosphere. *J. Geophys. Res.* **88**, 1145–1152 (1983).
- Dean, S. M., Minshall, T. A., Whitmarsh, R. B. & Louden, K. Deep structure of the ocean-continent transition in the southern Iberia Abyssal Plain from seismic refraction profiles: II The IAM-9 transect at 40° 20' N. *J. Geophys. Res.* **105**, 5859–5886 (2000).
- Chian, D., Louden, K. E., Minshall, T. A. & Whitmarsh, R. B. Deep structure of the ocean-continent transition in the southern Iberia Abyssal Plain from seismic refraction profiles: Ocean Drilling Program (Legs 149 and 173) transect. *J. Geophys. Res.* **104**, 7443–7462 (1999).
- Pickup, S. L. B., Whitmarsh, R. B., Fowler, C. M. R. & Reston, T. J. Insight into the nature of the ocean-continent transition of West Iberia from a deep multichannel seismic reflection profile. *Geology* **24**, 1079–1082 (1996).
- Whitmarsh, R. B. et al. in *Non-Volcanic Rifting of Continental Margins: A Comparison of Evidence from Land and Sea* (eds Wilson, R. C. L., Whitmarsh, R. B., Taylor, B. & Froitzheim, N.) 107–124 (Geological Society of London, London, in the press).
- Whitmarsh, R. B. & Wallace, P. J. in *Proc. ODP Sci. Results* (eds Beslier, M. O., Whitmarsh, R. B., Wallace, P. J. & Girardeau, J.) 1–36 (Ocean Drilling Program, College Station, Texas, 2001).
- Lagabrielle, Y. & Lemoine, M. Alpine, Corsican and Apennine ophiolites: the slow-spreading ridge model. *C. R. Acad. Sci. Ser. 2a* **325**, 909–920 (1997).
- Manatschal, G. & Bernoulli, D. Architecture and tectonic evolution of nonvolcanic margins: present-day Galicia and ancient Adria. *Tectonics* **18**, 1099–1199 (1999).
- Manatschal, G., Froitzheim, N., Rubenach, M. J. & Turrin, B. in *Non-Volcanic Rifting of Continental Margins: A Comparison of Evidence from Land and Sea* (eds Wilson, R. C. L., Whitmarsh, R. B., Taylor, B. & Froitzheim, N.) 405–428 (Geological Society of London, London, in the press).
- Manatschal, G. & Nievergelt, P. A continent-ocean transition recorded in the Err and Platta nappes (Eastern Switzerland). *Eclogae Geol. Helvetiae* **90**, 3–27 (1997).
- Froitzheim, N. & Manatschal, G. Kinematics of Jurassic rifting, mantle exhumation, and passive margin formation in the Austroalpine and Penninic nappes (eastern Switzerland). *Geol. Soc. Am. Bull.* **108**, 1120–1133 (1996).
- Desmurs, L., Manatschal, G. & Bernoulli, D. in *Non-Volcanic Rifting of Continental Margins: A Comparison of Evidence from Land and Sea* (eds Wilson, R. C. L., Whitmarsh, R. B., Taylor, B. & Froitzheim, N.) 235–266 (Geological Society of London, London, in the press).
- Boillot, G., Féraud, G., Recq, M. & Girardeau, J. Undercrusting by serpentinite beneath rifted margins. *Nature* **341**, 523–525 (1989).
- Reston, T. J., Krawczyk, C. M. & Kläschen, D. The S reflector west of Galicia (Spain). Evidence for detachment faulting during continental breakup from prestack depth migration. *J. Geophys. Res.* **101**, 8075–8091 (1996).
- Whitmarsh, R. B., Dean, S. M., Minshall, T. A. & Tompkins, M. Tectonic implications of exposure of lower continental crust beneath the Iberia Abyssal Plain, northeast Atlantic Ocean: geophysical evidence. *Tectonics* **19**, 919–942 (2000).
- Hermann, J., Müntener, O., Trommsdorff, V., Hansmann, W. & Piccardo, G. B. Fossil crust-to-mantle transition, Val Malenco (Italian Alps). *J. Geophys. Res.* **102**, 20123–20132 (1997).
- Müntener, O., Hermann, J. & Trommsdorff, V. Cooling history and exhumation of lower-crustal granulite and upper mantle (Malenco, eastern central Alps). *J. Petrol.* **41**, 175–200 (2000).
- Bertotti, G., Picotti, V., Bernoulli, D. & Casterlarin, A. From rifting to drifting tectonic evolution of the South-Alpine upper crust from the Triassic to the Early Cretaceous. *Sedim. Geol.* **86**, 53–76 (1993).
- Minshall, T. A., Dean, S. M., White, R. S. & Whitmarsh, R. B. in *Non-Volcanic Rifting of Continental Margins: A Comparison of Evidence from Land and Sea* (eds Wilson, R. C. L., Whitmarsh, R. B., Taylor, B. & Froitzheim, N.) 537–550 (Geological Society of London, London, in the press).
- Hébert, R., Gueddari, K., Laflèche, M. R., Beslier, M.-O. & Gardien, V. in *Non-Volcanic Rifting of Continental Margins: A Comparison of Evidence from Land and Sea* (eds Wilson, R. C. L., Whitmarsh, R. B., Taylor, B. & Froitzheim, N.) 161–189 (Geological Society of London, London, in the press).
- Abe, N. in *Non-Volcanic Rifting of Continental Margins: A Comparison of Evidence from Land and Sea* (eds Wilson, R. C. L., Whitmarsh, R. B., Taylor, B. & Froitzheim, N.) 143–159 (Geological Society of London, London, in the press).
- Skelton, A. D. L. & Valley, J. W. The relative timing of serpentinization and mantle exhumation at the ocean-continent transition, Iberia: constraints from oxygen isotopes. *Earth Planet. Sci. Lett.* **178**, 327–338 (2000).
- Schärer, U., Girardeau, J., Cornen, G. & Boillot, G. 138–121 Ma asthenospheric magmatism prior to continental break-up in the North Atlantic and geodynamic implications. *Earth Planet. Sci. Lett.* **181**, 555–572 (2000).
- Rampone, E. et al. Petrology, mineral and isotope geochemistry of the External Liguride peridotites (Northern Apennines, Italy). *J. Petrol.* **36**, 81–105 (1995).
- Müntener, O. & Hermann, J. in *Non-Volcanic Rifting of Continental Margins: A Comparison of Evidence from Land and Sea* (eds Wilson, R. C. L., Whitmarsh, R. B., Taylor, B. & Froitzheim, N.) 267–288 (Geological Society of London, London, in the press).
- Whitmarsh, R. B., Miles, P. R., Sibuet, J.-C. & Louvel, V. in *Proc. ODP Sci. Results* (eds Whitmarsh, R. B., Sawyer, D. S., Klaus, A. & Masson, D. G.) 665–674 (Ocean Drilling Program, College Station, Texas, 1996).

28. Harry, D. L. & Sawyer, D. S. A dynamic model of extension in the Baltimore Canyon Trough region. *Tectonics* **11**, 420–436 (1992).
29. Loudon, K. E. & Chian, D. in *Response of the Earth's Lithosphere to Extension* (eds White, R. S., Hardman, R. F. P., Watts, A. B. & Whitmarsh, R. B.) 767–799 (Phil. Trans. R. Soc. Ser. A, Royal Society, London, 1999).
30. Taylor, B., Goodliffe, A. M. & Martinez, F. How continents break up: Insights from Papua New Guinea. *J. Geophys. Res.* **104**, 7497–7512 (1999).

## Acknowledgements

We thank the Shipboard Scientific Parties of ODP Leg 149 and Leg 173 and those aboard RRS *Discovery* cruise 215. We thank the UK Natural Environment Research Council, The Royal Society of London and the Swiss National Science Foundation for support.

Correspondence and requests for materials should be addressed to R.B.W. (e-mail: rbw@soc.soton.ac.uk).

# Mitochondrial protein phylogeny joins myriapods with chelicerates

Ui Wook Hwang<sup>\*,†</sup>, Markus Friedrich<sup>‡</sup>, Diethard Tautz<sup>§</sup>, Chan Jong Park<sup>†</sup> & Won Kim<sup>†</sup>

<sup>\*</sup> Department of Biology, Teachers College, Kyungpook National University, Taegu 702-701, Korea

<sup>†</sup> School of Biological Sciences, Seoul National University, Seoul 151-742, Korea

<sup>‡</sup> Department of Biological Sciences, Wayne State University, 5047 Gullen Mall, Detroit, Michigan 48202, USA

<sup>§</sup> Abteilung für Evolutionsgenetik, Institut für Genetik, Universität zu Köln, Weyertal 121, 50931 Köln, Germany

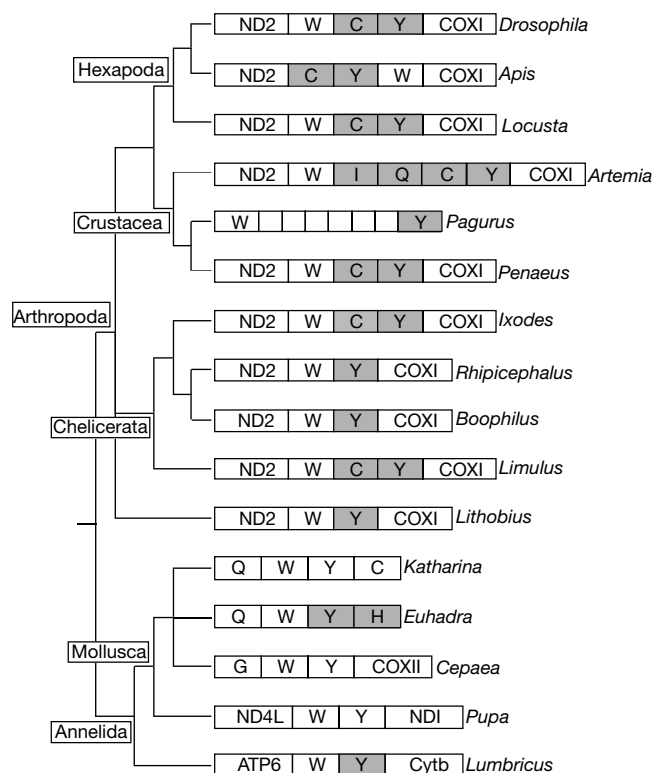
The animal phylum Arthropoda is very useful for the study of body plan evolution given its abundance of morphologically diverse species and our profound understanding of *Drosophila* development<sup>1</sup>. However, there is a lack of consistently resolved phylogenetic relationships between the four extant arthropod subphyla, Hexapoda, Myriapoda, Chelicerata and Crustacea. Recent molecular studies<sup>2–4</sup> have strongly supported a sister group relationship between Hexapoda and Crustacea, but have not resolved the phylogenetic position of Chelicerata and Myriapoda. Here we sequence the mitochondrial genome of the centipede species *Lithobius forficatus* and investigate its phylogenetic information content. Molecular phylogenetic analysis of conserved regions from the arthropod mitochondrial proteome yields highly resolved and congruent trees. We also find that a sister group relationship between Myriapoda and Chelicerata is strongly supported. We propose a model to explain the apparently parallel evolution of similar head morphologies in insects and myriapods.

The basal diversification of arthropod lineages, which date back into the late Cambrian period is still unclear. Morphological analyses<sup>5,6</sup> all suggest a monophyletic Arthropoda within which insects and myriapods are most closely related. Controversy, however, continued over whether insects, myriapods and crustaceans form a second major subclade, Mandibulata, on the basis of the shared derived possession of mandibles<sup>5</sup> or whether crustaceans are a sister group to chelicerates on the basis of the occurrence of biramous appendages in representatives of both groups<sup>6</sup>. Several independent molecular studies provided strong support for arthropod monophyly, a monophyletic Hexapoda, Myriapoda and Chelicerata, and, most significantly, a sister group relationship between insects and crustaceans (Pancrustacea) (for a review see ref. 7). Although they ruled out the possibility of insect/myriapod or crustacean/chelicerate sister clades, previous molecular studies did not resolve relationships between myriapods, chelicerates and

Pancrustacea<sup>2–4</sup>. Mitochondrial gene order rearrangements were initially interpreted to support a monophyletic Mandibulata<sup>8</sup>, but were later re-interpreted to further corroborate the Pancrustacea clade<sup>2</sup>.

Complete mitochondrial genome sequences can be informative at deep phylogenetic levels<sup>9</sup>. We therefore investigated their potential use for arthropod phylogeny. As examples of mitochondrial genomes are known from all arthropod subphyla except myriapods, we determined the complete mitochondrial genome sequence of the centipede *Lithobius forficatus*. The *Lithobius* mitochondrial genome is 15,437 base pairs (bp) (details will be given elsewhere). Gene content and arrangement correspond to that of conservatively evolving arthropod mitochondrial genomes with two exceptions. Most crustacean and insect mitochondrial genomes differ from *Lithobius* with regard to the position of the transfer RNA<sup>Leu(UUR)</sup> gene, which in crustaceans is located between the COXI and COXII genes and in *Lithobius* between the tRNA<sup>Leu(CUN)</sup> and ND1 genes. This is consistent with the previous demonstration that the COXI/tRNA<sup>Leu(UUR)</sup>/COXII arrangement is a synapomorphy of the Pancrustacea<sup>2</sup>.

Another difference concerns the position of the tRNA<sup>Cys</sup> gene, which in most arthropods resides between tRNA<sup>Trp</sup> and tRNA<sup>Tyr</sup> (Fig. 1), but in *Lithobius* it lies within the non-coding region of the



**Figure 1** Phylogenetic distribution of tRNA<sup>Cys</sup> arrangements in arthropod mitochondrial genomes. The relative location of tRNA<sup>Trp</sup> (W), tRNA<sup>Cys</sup> (C) and tRNA<sup>Tyr</sup> (Y) is shown for representative arthropod and outgroup species with similar arrangements. Multiple coding units separating tRNA<sup>Trp</sup> and tRNA<sup>Tyr</sup> in *Pagurus* are indicated by boxes. Transcription units in clear boxes code from left to right, those in shaded boxes code from right to left. The mollusc *Euhadra herklotsi* is the only non-arthropod species known so far in which tRNA<sup>Trp</sup> and tRNA<sup>Tyr</sup> are neighbours in opposite coding orientation, as in *Lithobius*. In a few non-arthropod species tRNA<sup>Trp</sup> and tRNA<sup>Tyr</sup> are next to each other, although in the same coding orientation. Re-examining non-annotated regions in published mitochondrial genome sequences, we found that the annelid species *Lumbricus terrestris* has coding probability for a second tRNA<sup>Tyr</sup>, which could result in a *Lithobius*-like tRNA<sup>Trp</sup> and tRNA<sup>Tyr</sup> arrangement (U.W.H., unpublished observation). This possibility, however, awaits confirmation by tRNA transcript analysis.



mitochondrial genome. Further exceptions are the honeybee *Apis mellifera*, the decapod *Pagurus longicarpus* and the tick species *Rhipicephalus sanguineus* and *Boophilus microplus*; these, however, represent lineages with exceptionally high rearrangement rates. In *Lithobius*, the tRNA<sup>Trp</sup> and tRNA<sup>Tyr</sup> genes lie directly next to each other in opposite coding directions, so we asked whether this or the tRNA<sup>Trp</sup>/tRNA<sup>Cys</sup>/tRNA<sup>Tyr</sup> arrangement is ancestral for arthropods. No non-arthropod mitochondrial genome is known to exhibit a tRNA<sup>Trp</sup>/tRNA<sup>Cys</sup>/tRNA<sup>Tyr</sup> arrangement, but the *Lithobius*-like tRNA<sup>Trp</sup>/tRNA<sup>Tyr</sup> arrangement also appears to be rare (Fig. 1)<sup>10</sup>. The inference of character state polarity is further confounded by considerable positional variation of the respective tRNAs across species. The tick species *Rhipicephalus sanguineus* and *Boophilus microplus* share the tRNA<sup>Trp</sup>/tRNA<sup>Tyr</sup> arrangement with *Lithobius*. This is due to parallel evolution, as indicated by the tRNA<sup>Trp</sup>/tRNA<sup>Cys</sup>/tRNA<sup>Tyr</sup> arrangement shared by the closely related tick species *Ixodes hexagonus* and the more distantly related horseshoe crab *Limulus polyphemus*. The tRNA<sup>Trp</sup>/tRNA<sup>Cys</sup>/tRNA<sup>Tyr</sup> arrangement is therefore with high certainty ancestral for chelicerates (Fig. 1). Thus, parallel evolution and the high evolutionary mobility of these particular tRNAs make the use of these tRNAs unreliable for deep-level cladistic analysis.

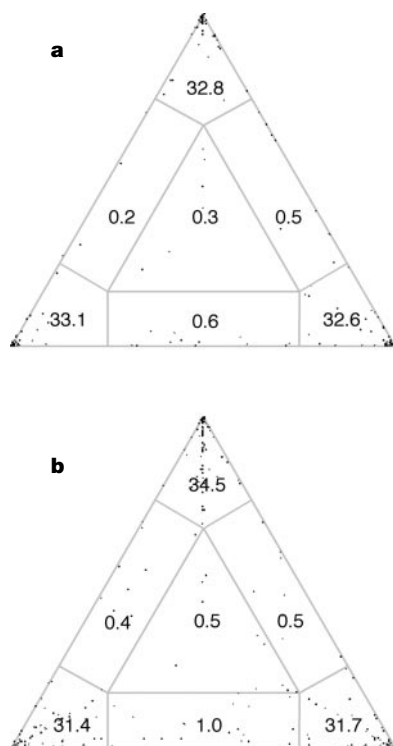
To explore the phylogenetic signal in mitochondrial protein sequences, we generated an alignment (18P2560) 2,560 amino acid sites long from conserved regions of 12 mitochondrial proteins from *Lithobius* and additional arthropod taxa. Annelid, mollusc and vertebrate species were added for outgroup comparison. Pairwise relative rate tests revealed that several species including the locust *Locusta migratoria*, the decapod *Pagurus longicarpus*, the

branchiopod species *Artemia franciscana*, the tick species *Ixodes hexagonus* and *Lithobius* exhibited significantly accelerated substitution rates. Furthermore, four species significantly departed from the average amino-acid composition in the alignment (Table 1). Nonetheless, maximum-likelihood mapping indicated a high phylogenetic information content in the alignment (Fig. 2a). Tree reconstruction with maximum-parsimony, distance and maximum-likelihood methods converged on a number of strongly supported clades (Fig. 3). Well established clades such as monophyletic Vertebrata, Eutrochozoa (*Lumbricus* and *Katharina*), Arthropoda, Decapoda (*Pagurus* and *Penaeus*) and Branchiopoda (*Artemia* and *Daphnia*), Chelicerata and Hexapoda were recovered with high branch-support values. Most basal nodes within arthropods were also consistently resolved. Decapods were strongly supported as a sister clade to insects, suggesting a paraphyletic Crustacea as recently noted<sup>11,12</sup>. Although the maximum-likelihood tree included a monophyletic Pancrustacea, branch-support analysis yielded little resolution with regard to the position of the Branchiopoda. The most striking result was a strong support for a sister group relationship between the myriapods and chelicerates with branch-support values equalling those of well established clades such as Chelicerata or Hexapoda.

To assess the impact of alignment site choice, we repeated tree estimation with alignments built from more stringently selected protein regions. The shortest alignment included 1,528 sites (18P1528), which exhibited an average maximum-likelihood distance two times lower than in the 18P2560 alignment, demonstrating considerable restriction to more slowly evolving sites (Table 1). This was associated with improved homogeneity of amino acid composition across taxa (Table 1). Maximum-likelihood mapping revealed a slight decrease in phylogenetic information content, which, in part, must be due to the reduction of sequence sample size (Fig. 2). Tree estimation yielded well resolved topologies, which were largely congruent with the results obtained with the 18P2560 alignment, the only difference being increased support for a monophyletic Pancrustacea (Fig. 3). These results suggest that the high resolution in the mitochondrial trees derives from the most slowly evolving protein regions. The consistent strong support for a monophyletic Myriapoda/Chelicerata demonstrates a robust phylogenetic signal for this clade in the mitochondrial proteins.

The support for a monophyletic Pancrustacea is conspicuously lower in the mitochondrial trees than in the nuclear ribosomal trees<sup>4</sup>, but the opposite applied to the support for the chelicerate/myriapod clade. This discrepancy is probably due to the combined effect of differences in taxon sampling and gene-specific fluctuations in the conservation of phylogenetic signal. Indeed, the support for a monophyletic Pancrustacea is much stronger in 28S than in the 18S nuclear ribosomal DNA sequences<sup>3</sup>.

It is essential to include closely related outgroup species to root the basal relationships of a phylogeny correctly. Recent studies suggest arthropods to be part of a higher clade, Ecdysozoa, of moulting animals including nematodes<sup>13</sup>. Although a monophyletic Ecdysozoa is not entirely consistently supported<sup>14,15</sup>, we considered the possibility that nematode mitochondrial sequences could

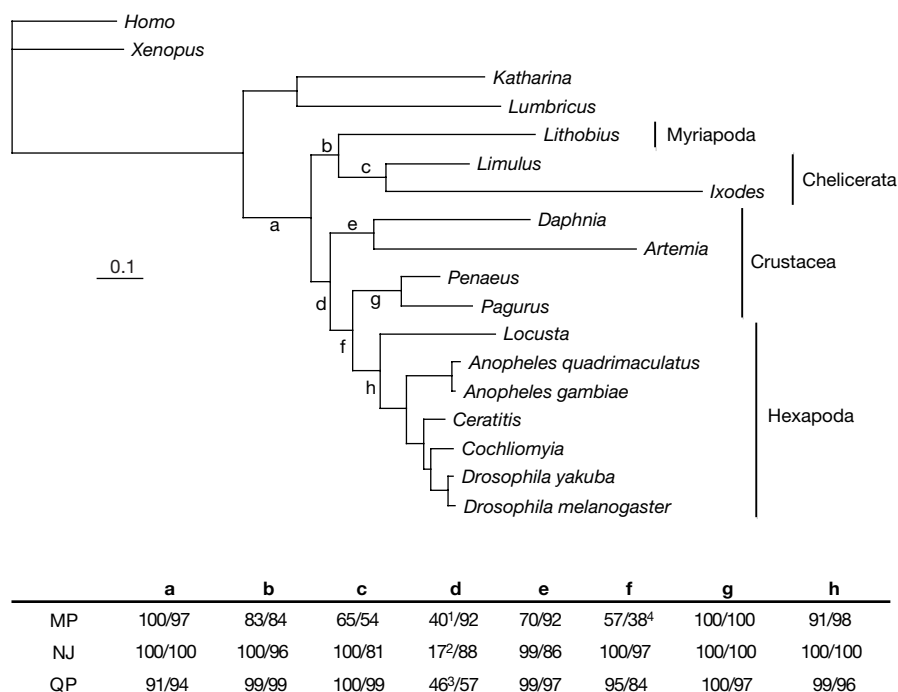


**Figure 2** Visualization of phylogenetic information content by maximum likelihood mapping. **a**, The 18P2560 alignment. **b**, The 18P1528 alignment. Maximum-likelihood trees for the total of 3,060 possible quartet combinations are mapped according to ref. 16. Corner regions contain the partition of fully resolved quartet trees, lateral regions contain the partition of partially resolved quartet trees, and the centre region contains the partition of completely unresolved quartet trees. For both alignments, more than 95% of all possible quartets are fully resolved, indicating high phylogenetic information content.

**Table 1** Comparison of multiple alignment features

|                                          | 18P2560                                                          | 18P1528       |
|------------------------------------------|------------------------------------------------------------------|---------------|
| Percentage of constant sites             | 31.7                                                             | 46.1          |
| $\alpha$                                 | 0.46                                                             | 0.39          |
| Average maximum-likelihood distance      | 0.77                                                             | 0.33          |
| Species with significant amino-acid bias | <i>Daphnia</i><br><i>Ixodes</i><br><i>Xenopus</i><br><i>Homo</i> | <i>Ixodes</i> |

Comparison of long (18P2560) and short (18P1528) multiple alignments with respect to percentage of constant sites, rate heterogeneity across sites as indicated by the  $\alpha$  parameter, average maximum-likelihood distance (substitutions per site) in pairwise species comparisons and partition of species with significantly deviating amino acid composition.



**Figure 3** Phylogram of best maximum-likelihood tree with 18P2560 alignment (ln(likelihood) = -42925.32). Bar represents 0.1 substitutions per site. Branches with letters have branch support values (BP) given below the tree for maximum parsimony (MP), neighbour-joining (NJ) and the maximum-likelihood-based quartet puzzling method (QP)<sup>28</sup>. Left numbers refer to 18P2560 alignment, right numbers to 18P1528 alignment.

Superscript numbers indicate branches that are not included in bootstrap majority rule consensus trees: 1, Branchiopoda placed at the base of arthropods with BP = 57; 2, Branchiopoda placed at the base of the arthropods with BP = 82; 3, Branchiopoda placed at the base of the arthropods with BP = 53; 4, monophyletic Crustacea supported with BP = 53.

represent a more adequate outgroup than the choice of the protostome and deuterostome sequences. Although nematode mitochondrial sequences are problematic for phylogeny reconstruction owing to dramatically accelerated substitution rates, a chelicerate/myriapod sister group clade was robust when nematodes were included in tree estimation (see the Supplementary Information). Nonetheless, it will be important to examine the effect of slowly evolving sequences from ecdysozoan taxa, particularly onychophorans, on the rooting of the arthropod mitochondrial tree.

A close link between myriapods and chelicerates has never, to our knowledge, been considered from a morphological perspective. We note, however, that the same grouping is tentatively supported in various analyses of nuclear ribosomal genes<sup>4,16,17</sup>. In addition, recent analyses of arthropod haemocyanin and *Hox* gene sequences point to a close relationship between chelicerates and myriapods<sup>18,19</sup>. Independent molecular data thus provide consistent support for a chelicerate/myriapod sister group relationship, arguing against a monophyletic Mandibulata. Future research is needed to examine the possibility of morphological synapomorphies for a chelicerate/myriapod clade.

Another important question is how similar head appendage arrays evolved in insects and myriapods, given the closer relationship of the latter to chelicerates. One possible scenario is that head segmentation and appendage differentiation in extant myriapods, insects and crustaceans is ancestral for arthropods. Chelicerate head morphology must then have evolved from a myriapod-like head morphology. Such evolutionary transformation is not inconceivable given that *Drosophila* head appendages have retained the potential to develop into primitive leg structures<sup>20</sup>. Nonetheless, the presence of largely undifferentiated postoral head appendages in primitive representatives of Trilobites and other extinct arthropods argues against this idea<sup>21</sup>. This is more consistent with the alternative possibility that the arthropod ancestor possessed a head with largely undifferentiated appendages from which myriapod and

insect head morphologies evolved in parallel. Recent comparisons of *Hox* gene expression revealed that arthropods share mechanisms of homeotic control of head segment specification even between groups as divergent as chelicerates and insects<sup>22,23</sup>. This implies that the myriapod and insect heads evolved from a common developmental grid despite an apparently more distant phylogenetic relationship. Their similarity may thus be the result of shared developmental constraints and parallel functional adaptation. Future comparative studies of arthropod head patterning should therefore reveal more similarities between the evolutionarily more closely related crustaceans and insects than myriapods. □

## Methods

### Sequence analysis

Total DNA was isolated from a specimen of the centipede species *Lithobius forficatus* collected in the garden of the Zoological Department of the University of München (Germany). A 538-bp portion of the large subunit ribosomal RNA gene (16S rDNA) was amplified by standard polymerase chain reaction (PCR) using universal primers (16SA: 5'-CGC CTG TTT ATC AAA AAC AT-3'; 16SB: 5'-CCG GTT GAA CTC AGA TCA-3') and sequenced. The complete genome was then amplified using the Expand Long Template PCR System (Roche Biochemicals) with the primers HPK16Saa (32mer), 5'-ATG CTA CCT TTG CAC RGT CAA GAT ACY GCG GC-3', and HPK16Sbb (34 mer), 5'-CTT ATC GAY AAA AAA GWT TGC GAC CTC GAT GTT G-3'. Cycling settings included one cycle of 2 min at 92 °C for initial denaturation, followed by 30 cycles of 10-s denaturation at 92 °C, 30-s annealing at 65 °C, and 13-min elongation at 68 °C. During the last 20 cycles, elongation times were increased for 20 s per cycle. The reaction was finished with a 20-min final elongation step at 68 °C. A single 15.5-kb-long PCR fragment was purified and used as a template for secondary PCR reactions. *EcoRI* or *XbaI* restriction fragments were cloned and sequenced in both directions on an ABI310 automated sequencer (Perkin Elmer). Overlaps between restriction fragment clones were confirmed by direct sequencing of PCR products spanning these regions. Protein-coding genes were identified by similarity of predicted amino-acid sequence with known mitochondrial protein sequences. The annotated sequence has been submitted to the EMBL data bank (accession number: AJ270997).

### Phylogenetic analysis

Complete mitochondrial genome sequences were retrieved from GenBank from the following arthropod species: fruitfly *Drosophila yakuba* (X03240), fruitfly *Drosophila*

*melanogaster* (U37541), mosquito *Anopheles quadrimaculatus* (L04272), mosquito *Anopheles gambiae* (L20934), medfly *Ceratitis capitata* (CCA242872), *Cochliomyia hominivorax* (AF260826), locust *Locusta migratoria* (X80245), honey bee *Apis mellifera* (L06178), brine shrimp *Artemia franciscana* (X69067), water flea *Daphnia pulex* (AF117817), shrimp *Penaeus monodon* (AF217843), hermit crab *Pagurus longicarpus* (AF150756), horseshoe crab *Limulus polyphemus* (AF216203), tick *Ixodes hexagonus* (AF081828), tick *Rhipicephalus sanguineus* (AF081829). For outgroup comparison, sequences were retrieved for the annelid *Lumbricus terrestris* (U24570), the mollusc *Katharina tunicata* (U09810), the nematodes *Caenorhabditis elegans* (X54252), *Ascaris suum* (X54253), *Trichinella spiralis* (AF293969) and *Onchocerca volvulus* (AF015193), and the vertebrate species *Homo sapiens* (J01415) and *Xenopus laevis* (M10217). Additional sequences were analysed for gene arrangements: *Boophilus microplus* (AF110613), *Euhadra herklotsi* (Z71696), *Cepaea nemoralis* (U23045) and *Pupa strigosa* (NC\_002176).

Multiple alignments were prepared for all putative protein sequences using Clustal W<sup>24</sup> at default settings. Consistent with previous studies<sup>25</sup>, preliminary analyses revealed obvious tree estimation artefacts due to extremely accelerated substitution rates or protein composition bias in the nematode species, the honeybee *Apis mellifera* and the tick species *Rhipicephalus sanguineus* and *Ixodes hexagonus*. With the exception of *Ixodes hexagonus* all of these taxa were therefore excluded from further analyses, reducing the total number of species considered to 18. Sequence alignment was repeated and inspected by eye for sufficient levels of sequence conservation, which resulted in the exclusion of the ATPase 8 gene (see Supplementary Information for single protein alignments). We used Gblocks<sup>26</sup> to extract regions of defined sequence conservation from the gene specific alignments and generate a single file of concatenated conserved regions. Default settings yielded the 18P2560 alignment. Modified parameter settings for generating the 18P1528 alignment were: minimum number of sequences for a conserved position: 15; maximum number of contiguous nonconserved positions: 2; minimum length of a block after gap cleaning: 5. Alignments can be retrieved from the EBI webserver (<ftp://ftp.ebi.ac.uk/pub/databases/emb/align>) under accession numbers ALIGN\_000111 and ALIGN\_000112. Maximum-likelihood mapping was carried out as described in ref. 16. Pairwise relative rate tests were carried out with the Hy-Phy program package<sup>27</sup>. Protein composition homogeneity test and maximum likelihood tree estimation was carried out using the TREE-PUZZLE program<sup>28</sup> applying the mtREV24 sequence evolution model for mitochondrial proteins<sup>29</sup> and a four rate approximated gamma distribution of among-site rate heterogeneity. Maximum-likelihood trees were determined by likelihood ratio tests between competing topologies. Maximum-parsimony tree reconstruction and neighbour-joining analysis with Dayhoff PAM matrix distances were performed using the respective algorithms implemented in Phylip 3.5 (ref. 30). Non-parametric bootstrapping analyses were based on 100 replicate data sets.

Received 22 February; accepted 13 July 2001.

1. Akam, M. Arthropods: developmental diversity within a (super) phylum. *Proc. Natl Acad. Sci. USA* **97**, 4438–4441 (2000).
2. Boore, J. L., Labrov, D. V. & Brown, W. M. Gene translocation links insects and crustaceans. *Nature* **392**, 667–668 (1998).
3. Shultz, J. W. & Regier, J. C. Phylogenetic analysis of arthropods using two nuclear protein-encoding genes supports a crustacean + hexapod clade. *Proc. R. Soc. Lond. B* **267**, 1011–1019 (2000).
4. Friedrich, M. & Tautz, D. Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods. *Nature* **376**, 165–167 (1995).
5. Snodgrass, R. E. Evolution of Annelida, Onychophora and Arthropoda. *Smithson. Misc. Coll.* **138**, 1–77 (1938).
6. Cline, J. L. Trilobites and the evolution of arthropods. *Science* **186**, 13–18 (1974).
7. Giribet, G. & Ribera, C. A review of arthropod phylogeny: new data based on ribosomal DNA sequences and direct character optimization. *Cladistics* **16**, 204–231 (2000).
8. Boore, J. L., Collins, T. M., Stanton, D., Daehler, L. L. & Brown, W. M. Deducing the pattern of arthropod phylogeny from mitochondrial DNA rearrangements. *Nature* **376**, 163–165 (1995).
9. Curule, J. P. & Kocher, T. D. Mitogenomics: digging deeper with complete mitochondrial genomes. *Trends Ecol. Evol.* **14**, 394–203 (1999).
10. Boore, J. L. Animal mitochondrial genomes. *Nucleic Acids Res.* **27**, 1767–1780 (1999).
11. Wilson, K., Cahill, V., Ballment, E. & Benzie, J. The complete sequence of the mitochondrial genome of the crustacean *Penaeus monodon*: are malacostracan crustaceans more closely related to insects than to branchiopods? *Mol. Biol. Evol.* **17**, 863–874 (2000).
12. Garcia-Machado, E. et al. Mitochondrial genes collectively suggest the paraphyly of Crustacea with respect to Insecta. *J. Mol. Evol.* **49**, 142–149 (1999).
13. Aguinaldo, A. et al. Evidence for a clade of nematodes, arthropods and other molting animals. *Nature* **387**, 489–493 (1997).
14. Sidow, A. & Thomas, W. K. A molecular evolutionary framework for eukaryotic model organisms. *Curr. Biol.* **4**, 596–603 (1994).
15. Hausdorf, B. Early evolution of the Bilateria. *Syst. Biol.* **49**, 130–142 (2000).
16. Strimmer, K. & von Haeseler, A. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proc. Natl Acad. Sci. USA* **94**, 6815–6819 (1997).
17. Peterson, K. J. & Eernisse, D. J. Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA sequences. *Evol. Dev.* **3**, 170–205 (2001).
18. Cook, C. E., Smith, M. L., Telford, M. J., Bastianello, A. & Akam, M. *Hox* genes and the phylogeny of the arthropods. *Curr. Biol.* **11**, 759–763 (2001).
19. Kusche, K. & Burmester, T. Diplopod hemocyanin sequence and the evolution of the Myriapoda. *Mol. Biol. Evol.* **18**, 1566–1573 (2001).
20. Casares, F. & Mann, R. S. Control of antennal versus leg development in *Drosophila*. *Nature* **392**, 723–726 (1998).
21. Delle Cave, L. & Simonetta, A. M. in *The Early Evolution of Metazoa and the Significance of Problematic Taxa* (eds Simonetta, A. M. & Conway Morris, S.) 189–244 (Cambridge Univ. Press, Cambridge, 1991).
22. Telford, M. J. & Thomas, R. H. Expression of homeobox genes shows chelicerate arthropods retain their deutocerebral segment. *Proc. Natl Acad. Sci. USA* **95**, 10671–10675 (1998).
23. Damen, W. G., Hausdorf, M., Seyfarth, E. A. & Tautz, D. A conserved mode of head segmentation in

arthropods revealed by the expression pattern of *Hox* genes in a spider. *Proc. Natl Acad. Sci. USA* **95**, 10665–10670 (1998).

24. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
25. Foster, P. G. & Hickey, D. A. Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. *J. Mol. Evol.* **48**, 284–290 (1999).
26. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).
27. Muse, S. V. & Kosakovsky Pond, S. L. *Hy-Phy 0.7 B* (North Carolina State Univ., Raleigh, 2000).
28. Strimmer, K. & von Haeseler, A. Quartet puzzling—a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**, 964–969 (1996).
29. Adachi, J. & Hasegawa, M. Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Mol. Evol.* **42**, 459–468 (1996).
30. Felsenstein, J. PHYLIP (Phylogeny Inference Package). (Univ. Washington, Seattle, 1995).

Supplementary information is available from Nature's World-Wide Website (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

## Acknowledgements

We thank N. Muqim for technical assistance and A. Minelli and T. Burmester for comments on the manuscript. Most computation was performed on the Biological Software Server of the Institute Pasteur Paris. This study was in part supported by a DFG grant to D.T. and a Brain Korea 21 Project grant to W.K. U.W.H. was supported by fellowships from Deutscher Akademischer Austauschdienst, Korea Science and Engineering Foundation, and the Brain Korea 21 Project (Medical Sciences, Yonsei University).

Correspondence and requests for materials should be addressed to U.W.H. (e-mail: [uwhwang@knu.ac.kr](mailto:uwhwang@knu.ac.kr)) or M.F. (e-mail: [mf@biology.biosci.wayne.edu](mailto:mf@biology.biosci.wayne.edu)).

# Arthropod phylogeny based on eight molecular loci and morphology

Gonzalo Giribet\*, Gregory D. Edgecombe† & Ward C. Wheeler‡

\* Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

† Australian Museum, 6 College Street, Sydney, New South Wales 2010, Australia

‡ Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024, USA

The interrelationships of major clades within the Arthropoda remain one of the most contentious issues in systematics, which has traditionally been the domain of morphologists<sup>1,2</sup>. A growing body of DNA sequences and other types of molecular data has revitalized study of arthropod phylogeny<sup>3–7</sup> and has inspired new considerations of character evolution<sup>8,9</sup>. Novel hypotheses such as a crustacean–hexapod affinity<sup>4,10–12</sup> were based on analyses of single or few genes and limited taxon sampling, but have received recent support from mitochondrial gene order<sup>13</sup>, and eye and brain ultrastructure and neurogenesis<sup>14,15</sup>. Here we assess relationships within Arthropoda based on a synthesis of all well sampled molecular loci together with a comprehensive data set of morphological, developmental, ultrastructural and gene-order characters. The molecular data include sequences of three nuclear ribosomal genes, three nuclear protein-coding genes, and two mitochondrial genes (one protein coding, one ribosomal). We devised new optimization procedures<sup>16,17</sup> and constructed a parallel computer cluster with 256 central processing units<sup>18</sup> to analyse molecular data on a scale not previously possible. The optimal 'total evidence' cladogram supports the crustacean–hexapod clade, recognizes pycnogonids as sister to other euarthropods, and indicates monophyly of Myriapoda and Mandibulata.

Based on morphological evidence, neontological<sup>1,5,6</sup> and palaeontological<sup>2</sup> hypotheses regarding deep divergences within Arthropoda differ in the monophyly of Mandibulata (arthropods with mandibles: crustaceans, myriapods and hexapods) versus



Schizoramia (arthropods with biramous appendages: crustaceans and stem-group chelicerates). Both fields have traditionally agreed that atelocerates (myriapods and hexapods) constitute a natural group. The Pancrustacea<sup>19</sup> hypothesis challenges a myriapod–hexapod alliance, instead resolving crustaceans as the closest relatives of hexapods. Only by combining all of the available data can the competing hypotheses be compared effectively for their ability to explain the evolution of the diverse sorts of changes that have occurred in many aspects of arthropod biology. Simultaneous analyses of morphological and molecular data addressed early conflicts of character data quality in arthropod evolution<sup>3,5,6,20</sup>. The investigation of a few more gene systems<sup>12,21,22</sup> complemented the widely used set of nuclear ribosomal genes<sup>5,7,10</sup>. However, to our knowledge, synthetic work combining multiple sources of molecular information, together with morphological data, has been restricted in either taxonomic or locus sampling.

We analysed sequence data of three nuclear ribosomal genes (18S ribosomal RNA, the D3 region of 28S rRNA, and the small nuclear rRNA U2), three nuclear protein-coding genes (histone H3, elongation factor-1 $\alpha$  and the largest subunit of RNA polymerase II), and

two mitochondrial loci, one protein coding (cytochrome *c* oxidase I) and one ribosomal (16S rRNA). From these data, six loci have been mostly sequenced in ours and co-workers' laboratories, and the remaining two genes (elongation factor-1 $\alpha$  and RNA polymerase II) were obtained from GenBank (see Table 1). The sequence data comprise about 5 kilobases (kb) of nucleotide information. We chose these genes because of their dense taxonomic sampling and because they have been reported to resolve phylogenetic history at different taxonomic levels, so we expected their levels of resolution to overlap.

Non-sequence data analyses for arthropod interrelationships are few, and all rely on groundplan coding, a strategy that tends to standardize states within 'known' morphologically defined groups. This coding strategy tends to generate phylogenetic trees in agreement with classical hypotheses, but represents a biased test for relationships. On the contrary, if character states are scored only for those taxa for which real observations have been made, missing data (unknown states) may proliferate, which more accurately represents the actual knowledge of the group. Following the logic of this coding strategy, we modified, expanded and reanalysed a published

**Table 1 Taxonomic categories represented in the trees, exemplar species and genes sequenced for each terminal**

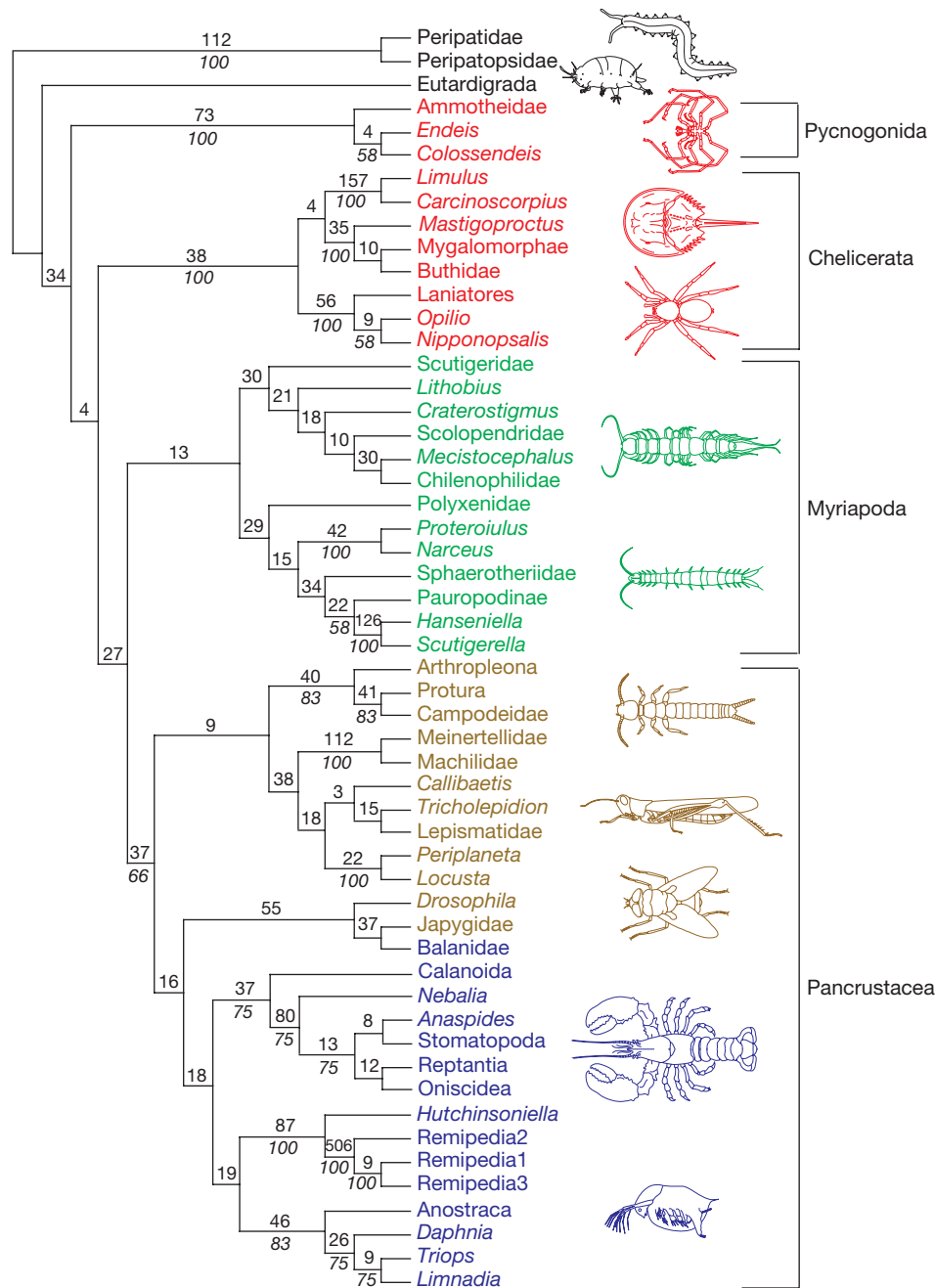
|                        |                                                     |     |     |    |    |    |     |     |     |
|------------------------|-----------------------------------------------------|-----|-----|----|----|----|-----|-----|-----|
| Peripatopsidae         | <i>Euperipatoides/Peripatopsis</i>                  | 18S | 28S | H3 | U2 | EF | POL | COI |     |
| Peripatidae            | <i>Epiperipatus/Oroperipatus</i>                    | 18S |     | H3 | U2 | EF | POL | COI |     |
| Eutardigrada           | <i>Macrobiotus/Milnesium</i>                        | 18S |     |    | U2 | EF | POL |     |     |
| MYCOGONIDA             |                                                     |     |     |    |    |    |     |     |     |
| <i>Endeis</i>          | <i>Endeis laevis</i>                                | 18S | 28S |    |    | EF | POL | COI |     |
| <i>Colossendeis</i>    | <i>Colossendeis</i> sp.                             | 18S | 28S |    |    | EF | POL | COI |     |
| Ammonoidea             | <i>Achelia/Ammonoella/Tanystylum</i>                | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| EUCHELICERATA          |                                                     |     |     |    |    |    |     |     |     |
| <i>Limulus</i>         | <i>Limulus polyphemus</i>                           | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| <i>Carcinoscorpius</i> | <i>Carcinoscorpius rotundicaudus</i>                | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| Buthidae               | <i>Androctonus/Lychas</i>                           | 18S | 28S | H3 | U2 |    |     | COI | 16S |
| <i>Mastigoproctus</i>  | <i>Mastigoproctus giganteus</i>                     | 18S | 28S |    |    | EF | POL | COI | 16S |
| Mygalomorphae          | <i>Atrax/Aphonopelma</i>                            | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| <i>Opilio</i>          | <i>Opilio parietinus</i>                            | 18S | 28S |    |    |    |     | COI | 16S |
| <i>Nipponopsalis</i>   | <i>Nipponopsalis abei</i>                           | 18S | 28S |    |    | EF | POL |     |     |
| Laniatores             | <i>Equitius/Vonones/Dalquestia</i>                  | 18S | 28S | H3 | U2 | EF | POL | COI |     |
| MYRIAPODA              |                                                     |     |     |    |    |    |     |     |     |
| Scutigeridae           | <i>Allothreua/Scutigera</i>                         | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| <i>Lithobius</i>       | <i>L. obscurus/variegatus</i>                       | 18S | 28S | H3 |    | EF | POL | COI | 16S |
| <i>Craterostigma</i>   | <i>Craterostigma tasmanianus</i>                    | 18S | 28S | H3 | U2 |    |     | COI | 16S |
| Scolopendridae         | <i>Cormocephalus/Scolopendra/Ethmostigma</i>        | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| <i>Mecistocephalus</i> | <i>Mecistocephalus</i> sp./tahiensis                | 18S | 28S | H3 |    |    |     | COI | 16S |
| Chilopodidae           | <i>Ribautia/Pachymyrmex</i>                         | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| <i>Hansenella</i>      | <i>Hansenella</i> sp.                               | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| <i>Scutigera</i>       | <i>Scutigera</i> sp.                                | 18S | 28S |    |    | EF | POL |     |     |
| Paupodinae             | <i>Paupodinae</i> sp.                               | 18S | 28S | H3 | U2 |    |     |     |     |
| Polyxenidae            | <i>Polyxenus/Unixenus</i>                           | 18S | 28S | H3 | U2 | EF | POL | COI |     |
| Sphaerotheriidae       | <i>Cyliosomatini/Epicyliosoma</i> sp.               | 18S |     | H3 | U2 |    |     | COI | 16S |
| <i>Proteroiulus</i>    | <i>Proteroiulus fuscus</i>                          | 18S | 28S |    |    | EF | POL | COI | 16S |
| <i>Narceus</i>         | <i>Narceus americanus</i>                           | 18S | 28S |    |    | EF | POL | COI | 16S |
| CRUSTACEA              |                                                     |     |     |    |    |    |     |     |     |
| Remipedia              | <i>Speleonectes/Lasionectes</i>                     | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| <i>Hutchinsoniella</i> | <i>Hutchinsoniella macracantha</i>                  | 18S | 28S | H3 |    | EF | POL | COI | 16S |
| Calanoida              | <i>Calanus/Eurytemora</i>                           | 18S |     |    |    | EF | POL | COI | 16S |
| Anostraca              | <i>Artemia/Branchinella</i>                         | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| <i>Triops</i>          | <i>T. longicaudatus/australiensis</i>               | 18S |     | H3 | U2 | EF | POL | COI | 16S |
| <i>Limnadia</i>        | <i>Limnadia</i> sp./lenticularis                    | 18S |     |    |    | EF | POL |     |     |
| <i>Daphnia</i>         | <i>Daphnia galeata/pulex</i> sp.                    | 18S |     |    | U2 |    |     | COI | 16S |
| <i>Nebalia</i>         | <i>Nebalia</i> sp./longicornis/hessleri             | 18S | 28S | H3 | U2 | EF | POL |     |     |
| Balanidae              | <i>Balanus/Semibalanus</i>                          | 18S | 28S | H3 | U2 | EF | POL | COI |     |
| Stomatopoda            | <i>Kempina mikado/Gonodactylus</i>                  | 18S |     | H3 | U2 |    |     | COI | 16S |
| <i>Anaspides</i>       | <i>Anaspides tasmaniae</i>                          | 18S | 28S |    |    |    |     | COI | 16S |
| Oniscidea              | <i>Armadillidium/Trichoniscus/Australophiloscia</i> | 18S | 28S |    |    | EF | POL | COI | 16S |
| Reptantia              | <i>Homarus americanus/Libinia</i>                   | 18S | 28S | H3 | U2 | EF |     | COI | 16S |
| HEXAPODA               |                                                     |     |     |    |    |    |     |     |     |
| Protura                | <i>Acerentulus/Nipponentomon/Acerentomon</i>        | 18S | 28S | H3 |    | EF |     |     |     |
| Arthropleona           | <i>Podura/Archisotoma/Tomocerus</i>                 | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| Campodeidae            | <i>Campodea/Eumesocampa</i>                         | 18S | 28S | H3 |    | EF | POL | COI | 16S |
| Japygidae              | <i>Cataglyphis/Metaglyphis/Heteroglyphis</i>        | 18S | 28S |    |    | EF | POL | COI | 16S |
| Meinertellidae         | <i>Allomachilis/Machiloides</i>                     | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |
| Machilidae             | <i>Dilta/Petrobiinae/Pedetontus</i>                 | 18S | 28S | H3 |    | EF | POL | COI | 16S |
| <i>Tricholepidion</i>  | <i>Tricholepidion gertschii</i>                     | 18S | 28S | H3 | U2 |    |     |     |     |
| Lepismatidae           | <i>Thermobia/Ctenolepisma</i>                       | 18S | 28S |    |    | EF | POL | COI | 16S |
| <i>Callibaetis</i>     | <i>Callibaetis ferrugineus</i>                      | 18S |     | H3 | U2 |    |     | COI | 16S |
| <i>Periplaneta</i>     | <i>Periplaneta americana</i>                        | 18S |     | H3 | U2 | EF | POL | COI |     |
| <i>Locusta</i>         | <i>Locusta migratoria</i>                           | 18S | 28S | H3 | U2 |    |     | COI | 16S |
| <i>Drosophila</i>      | <i>Drosophila melanogaster</i>                      | 18S | 28S | H3 | U2 | EF | POL | COI | 16S |

The taxonomic category listed is the least inclusive for the species for which molecular data have been obtained. More details on exact GenBank numbers for each partition, fragments analysed and batch files for the analyses is provided as Supplementary Information.

morphology data set<sup>6</sup> in order to code the non-sequence (morphological and gene order) features of 51 terminals (2 onychophoran, 1 tardigrade and 48 arthropod lineages) for which at least four of the eight selected loci were available. A total of 303 characters was employed (see Supplementary Information).

The principle of total evidence is an important maxim in phylogenetic systematics because alternatives to using all available evidence must explicitly exclude parts of the data<sup>23</sup>. Excluding data *a priori* biases the outcome towards the set of characters retained. Combining all sources of molecular information with non-sequence characters allows phylogenetic hypotheses to be formulated based on a maximum number of independent data points and thus provides the greatest explanatory power.

The shortest cladograms based on the morphology and gene-order data alone recognize Mandibulata, Crustacea and Hexapoda as well supported clades (Bremer support  $\geq 6$ ). A sister-group relationship between Hexapoda and monophyletic Myriapoda is weakly favoured over the alternative grouping of Crustacea and Hexapoda, one step separating these hypotheses. The tree resulting from the combined analysis of all sources of data for the parameter set that minimizes overall incongruence among data partitions (using a character-based congruence metric as an optimality criterion<sup>24</sup>) shows monophyly of arthropods (Fig. 1). Pycnogonids are a sister group to all other euarthropods, as is also resolved based on the non-sequence data alone. Chelicerata (horseshoe crabs and arachnids) and Mandibulata are sister taxa within the



**Figure 1** Phylogenetic tree of arthropod lineages based on DNA sequence data of eight loci and 303 characters of non-sequence data for the parameter set that minimizes overall incongruence among the nine partitions. The single tree, obtained when insertions and deletions equal all other genetic changes, requires 27,375 steps. Bremer support

values are shown above branches while percentage of analytical parameters yielding a given topology are represented in italics. 'Chelicerates', myriapods, hexapods and crustaceans are displayed in red, green, brown and blue, respectively.

Euarthropoda. The data support neither a monophyletic Schizoramia (chelicerates + crustaceans) nor a clade composed exclusively of chelicerates and myriapods.

A monophyletic group is formed by myriapods, crustaceans and hexapods, a resolution that is consistent with the origin of mandibles being a unique event during arthropod evolution<sup>8,9</sup>. Relationships within Mandibulata strongly support the monophyly of Pancrustacea, although support is weak for the monophyly of myriapods. The basic division of Myriapoda into Chilopoda (centipedes) and Progoneata (millipedes, pauropods and symphylans) is retrieved by morphology as well as the combination of all data. Resolution of the internal phylogeny of the centipedes is identical between morphological and total-evidence trees, whereas relationships of the progoneates differ substantially. Inclusion of sequence data serves to unite the Symphyla and Pauropoda, which perturbs the monophyly of Diplopoda (millipedes).

Pancrustacean relationships are confusing in the optimal cladogram based on combined data, because neither crustaceans nor hexapods are strictly monophyletic. *Drosophila* forms a clade with a japygid dipluran (basal hexapod) and a barnacle, and this clade receives a moderately high Bremer support value. However, a clade resolving *Drosophila* with other insects, as expected from the morphological analysis, is found in 91% of the combined analyses (all other parameter sets explored). The aberrant behaviour of *Drosophila* is mainly due to its placement based on the gene fragments of elongation factor-1 $\alpha$  and RNA polymerase II. Unfortunately *Drosophila* was not included in previous analyses with these gene fragments.

Crustacea is monophyletic in the non-sequence data analysis as well as in 50% of the combined analyses performed for the different parameter sets. Under the remaining six analytical parameters, Balanidae (and in one instance *Daphnia*) do not group with the remaining crustaceans. This result is more readily interpreted as anomalous behaviour of the sequence data of one organism than as evidence for crustacean paraphyly, as the original Pancrustacea hypothesis proposed. Morphological and combined analyses both identify the major groups Malacostraca and Branchiopoda, with the same basal splits (Leptostraca + Eumalacostraca and Anostraca + Phyllopoda, respectively) present in each cladogram. One of the most strongly supported groups within Crustacea, a union of cephalocarids and remipedes, is unexpected morphologically but has been detected in previous sequence analyses<sup>25</sup>.

The analysis of deep evolutionary relationships has never been easy, especially when we attempt to make global statements about the most diverse metazoan phylum using a limited sample of creatures and characters. To our knowledge, this is the largest data set compiled for such an important group of animals and was analysed using unique cluster computer technology to provide a comprehensive hypothesis of arthropod relationships. Both quality and quantity of data were optimized to create precise and robust hypotheses. Such an approach is the first step towards comprehensive genomic comparisons and the understanding of the animal group that has dominated for the past 500 million years. □

## Methods

### Non-sequence data

Morphological and gene-order characters were analysed with NONA v. 2.0 (ref. 26) using a standard 1,000 replicates of random addition sequence, followed by TBR (tree bisection and reconnection) branch swapping. Multistate characters were unordered except where justified in character discussions.

### Direct optimization

Phylogenetic analysis was performed by combining all of the available evidence and searching for the simplest (most parsimonious) explanation for character variation. This was accomplished with the program POY (ref. 17), which is explicitly designed to reconstruct phylogeny from diverse sources of information. Molecular data were optimized using the 'direct optimization' procedure<sup>16</sup>, which derives cladogram costs without multiple sequence alignment. This methodology accommodates sequence length variation as transformations involving the addition, deletion and substitution of

nucleotides, as opposed to the conjuring of unobservable 'gaps'. Direct optimization produces more parsimonious and more congruent results than multiple sequence alignment<sup>27</sup>, and in a vastly accelerated time frame. POY minimizes the weighted number of evolutionary changes over the entire tree, working in a one-step fashion as opposed to the more classical two-step analyses (alignment + tree search).

### Node support

Bremer support<sup>28</sup> was used as a measure for node support. This measure records the difference in length between the favoured tree and the shortest tree that does not contain a given group. As such, the Bremer value is a measure of how decisive the data are for a given branch of the tree.

### Sensitivity analysis

To test nodal stability, 12 sets of parameters (gap/change quotients of 1, 2, 4 and transversion/transition quotients of 1, 2, 4 and  $\infty$ ) were analysed for each of the ten partitions (eight loci, the combined molecular data, and the combined molecular + morphological data). Each of these 120 independent analyses was executed in parallel in the 256 processors, totalling 2 months of intense computation time using extremely effective tree search algorithms<sup>29</sup> and an aggressive search strategy, equivalent to 42 years of computing time if analyses had to be conducted in a single-processor machine. Parallelization was executed in groups of 32 processors, the point of maximum efficiency for marshalling jobs in our cluster. This kind of analysis explores the stability of data to analysis parameters, and therefore allows hypotheses to be formulated in a more robust way than traditional phylogenetic analysis.

The simultaneous analysis of multiple genes, morphology and mitochondrial gene order hinders application of different evolutionary models to independent partitions, and even more so when multiple parameters are explored. Therefore we opted for the simplicity of treating all partitions equally. Similarly, morphological and gene-order changes were weighted equally with insertion and deletion events.

Our 'optimal tree' is that which minimizes tree length (the most parsimonious) for all the data analysed in combination (total evidence) and for the parameter set that minimizes overall incongruence among partitions<sup>24</sup>. In this case overall congruence is calculated as per the incongruence length difference (ILD) test<sup>30</sup>, and it is optimized at an insertion/deletion cost equal to all other transformations (transversions or transitions).

### Search strategy

Tree search commands executed in POY included random addition sequence followed by a fast parallelized tree-building step and by SPR (subtree pruning and regrafting) and TBR branch swapping. When classical swapping algorithms did not improve tree length, the data were submitted to several rounds of tree drifting and tree fusing<sup>29</sup> to decrease tree length. The entire search strategy was repeated up to 1,000 times or until the results converged on the same result at least three times in independent replicates. A typical command line used for a given stepmatrix (-molecularmatrix 111) is as follows:

```
poy -parallel -jobspernode 2 -noleading -norandomizeoutgroup
-molecularmatrix 111 -multibuild 10 -buildspr -buildtbr
-approxbuild -buildmaxtrees 1 -random 1000 -stopat 3 -minstop 10
-multirandom -sprmaxtrees 1 -tbrmaxtrees 2 -numdriftchanges 30
-driftspr -numdriftspr 10 -drifttbr -numdrifttbr 10 -0tchtrees
-holdmaxtrees 100 -controllers 32 -maxtrees 20 -treefuse
-fuselimit 10 -fusingroup 5 -seed -1 'input 0les'
-prealigned 'input protein-coding 0les' > tot111.out 2>
tot111.err
```

The specific command lines and data files are provided as Supplementary Information.

Received 31 January; accepted 31 July 2001.

- Snodgrass, R. E. Evolution of the Annelida, Onychophora and Arthropoda. *Smithson. Misc. Collect.* **97**, 1–159 (1938).
- Wills, M. A., Briggs, D. E. G., Fortey, R. A., Wilkinson, M. & Sneath, P. H. A. *Arthropod Fossils and Phylogeny* (ed. Edgecombe, G. D.) 33–105 (Columbia Univ. Press, New York, 1998).
- Wheeler, W. C., Cartwright, P. & Hayashi, C. Y. Arthropod phylogeny: A combined approach. *Cladistics* **9**, 1–39 (1993).
- Regier, J. C. & Shultz, J. W. Molecular phylogeny of arthropods and the significance of the Cambrian "explosion" for molecular systematics. *Am. Zool.* **38**, 918–928 (1998).
- Wheeler, W. C. *Arthropod Fossils and Phylogeny* (ed. Edgecombe, G. D.) 9–32 (Columbia Univ. Press, New York, 1998).
- Edgecombe, G. D., Wilson, G. D. E., Colgan, D. J., Gray, M. R. & Cassis, G. Arthropod cladistics: Combined analysis of histone H3 and U2 snRNA sequences and morphology. *Cladistics* **16**, 155–203 (2000).
- Giribet, G. & Ribera, C. A review of arthropod phylogeny: New data based on ribosomal DNA sequences and direct character optimization. *Cladistics* **16**, 204–231 (2000).
- Popadić, A., Panganiban, G., Rusch, D., Shear, W. A. & Kaufman, T. C. Molecular evidence for the gnathobasic derivation of arthropod mandibles and for the appendicular origin of the labrum and other structures. *Dev. Genes Evol.* **208**, 142–150 (1998).
- Scholtz, G., Mittmann, B. & Gerberding, M. The pattern of *Distal-less* expression in the mouthparts of crustaceans, myriapods and insects: new evidence for a gnathobasic mandible and the common origin of Mandibulata. *Int. J. Dev. Biol.* **42**, 801–810 (1998).
- Friedrich, M. & Tautz, D. Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods. *Nature* **376**, 165–167 (1995).
- Giribet, G., Carranza, S., Bagnuà, J., Riutort, M. & Ribera, C. First molecular evidence for the existence of a Tardigrada + Arthropoda clade. *Mol. Biol. Evol.* **13**, 76–84 (1996).



12. Shultz, J. W. & Regier, J. C. Phylogenetic analysis of arthropods using two nuclear protein-encoding genes supports a crustacean + hexapod clade. *Proc. R. Soc. Lond. B* **267**, 1011–1019 (2000).
13. Boore, J. L., Lavrow, D. V. & Brown, W. M. Gene translocation links insects and crustaceans. *Nature* **392**, 667–668 (1998).
14. Dohle, W. Are the insects terrestrial crustaceans? A discussion of some new facts and arguments and the proposal of the proper name "Tetraconata" for the monophyletic unit Crustacea + Hexapoda. *Ann. Soc. Entomol. France* **37**, 85–103 (2001).
15. Strausfeld, N. J. Crustacean–insect relationships: The use of brain characters to derive phylogeny amongst segmented invertebrates. *Brain Behav. Evol.* **52**, 186–206 (1998).
16. Wheeler, W. C. Optimization alignment: The end of multiple sequence alignment in phylogenetics? *Cladistics* **12**, 1–9 (1996).
17. Wheeler, W. C. & Gladstein, D. POY (2000). Program and documentation available at [ftp://amnh.org/pub/molecular](http://amnh.org/pub/molecular).
18. Gee, H. Homegrown computer roots out phylogenetic networks. *Nature* **404**, 214 (2000).
19. Zrzavý, J. & Štys, P. The basic body plan of arthropods: Insights from evolutionary morphology and developmental biology. *J. Evol. Biol.* **10**, 353–367 (1997).
20. Zrzavý, J., Hypša, V. & Vlásková, M. *Arthropod Relationships* (eds Fortey, R. A. & Thomas, R. H.) 97–107 (Chapman & Hall, London, 1998).
21. Regier, J. C. & Shultz, J. W. Molecular phylogeny of the major arthropod groups indicates polyphyly of crustaceans and a new hypothesis for the origin of hexapods. *Mol. Biol. Evol.* **14**, 902–913 (1997).
22. Colgan, D. J. *et al.* Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Aust. J. Zool.* **46**, 419–437 (1998).
23. Kluge, A. G. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Syst. Zool.* **38**, 7–25 (1989).
24. Wheeler, W. C. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. *Syst. Biol.* **44**, 321–331 (1995).
25. Spears, T. & Abele, L. G. *Arthropod Relationships* (eds Fortey, R. A. & Thomas, R. H.) 169–187 (Chapman & Hall, London, 1998).
26. Goloboff, P. A. NONA v. 2.0 (1998). Program and documentation available at [ftp://unt.edu.ar/pub/parsimony](http://unt.edu.ar/pub/parsimony).
27. Wahlberg, N. & Zimmermann, M. Pattern of phylogenetic relationships among members of the tribe Melitaeini (Lepidoptera: Nymphalidae). *Cladistics* **16**, 347–363 (2000).
28. Bremer, K. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* **42**, 795–803 (1988).
29. Goloboff, P. A. Analyzing large data sets in reasonable times: Solutions for composite optima. *Cladistics* **15**, 415–428 (1999).
30. Farris, J. S., Källersjö, M., Kluge, A. G. & Bult, C. Constructing a significance test for incongruence. *Syst. Biol.* **44**, 570–572 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

#### Acknowledgements

We thank all those who have assisted us with morphological discussions, laboratory work (especially K. Demeo), collecting specimens, and given any other form of help or advice. D. Colgan and G. Wilson have been valued collaborators. S. Thurston provided technical illustration. J. Shultz and L. Prendini shared specimens and unpublished sequence data. Funding was mainly provided by the Fundamental Biology Program of NASA.

Correspondence and requests for materials should be addressed to G.G. (e-mail: [ggribet@oeb.harvard.edu](mailto:ggribet@oeb.harvard.edu)). GenBank accession codes for the new sequences are AF370781–AF370876.

## Dissociation between hand motion and population vectors from neural activity in motor cortex

Stephen H. Scott, Paul L. Gribble\*, Kirsten M. Graham & D. William Cabel

CIHR Group in Sensory-Motor Systems, Centre for Neuroscience Studies, Department of Anatomy and Cell Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada

The population vector hypothesis was introduced almost twenty years ago to illustrate that a population vector constructed from neural activity in primary motor cortex (MI) of non-human primates could predict the direction of hand movement during reaching<sup>1–6</sup>. Alternative explanations for this population signal

have been suggested<sup>7,8</sup> but could not be tested experimentally owing to movement complexity in the standard reaching model. We re-examined this issue by recording the activity of neurons in contralateral MI of monkeys while they made reaching movements with their right arms oriented in the horizontal plane—where the mechanics of limb motion are measurable and anisotropic. Here we found systematic biases between the population vector and the direction of hand movement. These errors were attributed to a non-uniform distribution of preferred directions of neurons and the non-uniformity covaried with peak joint power at the shoulder and elbow. These observations contradict the population vector hypothesis and show that non-human primates are capable of generating reaching movements to spatial targets even though population vectors based on MI activity do not point in the direction of hand motion.

Primary motor cortex (MI) has an important function in controlling visually guided limb movements, and a central problem in motor research has been to identify how neurons within MI participate in these tasks<sup>4,9</sup>. Several studies have shown that the activity of individual neurons is sensitive to many different parameters related to target, hand and limb movement<sup>5,10–13</sup>. However, it has been shown that if individual cells are represented as vectors, with direction defined by the cell's preferred direction (PD), the direction of movement in which the cell is maximally active) and magnitude defined as the cell's discharge rate for a given movement direction, the resulting vector sum (population vector) of all cell vectors is congruent with the direction of hand movement<sup>1–3,6</sup>. This ability to predict hand motion has supported the idea that MI may reflect a higher level representation related to movement direction<sup>4,5</sup>. Theoretical studies, however, have argued that neural activity in myriad coordinate frames related to sensory or motor features of the task would also produce population vectors that point in the direction of hand motion<sup>7,8,14</sup>. This debate on the neural representation of movement in MI is important not only for understanding its role in movement planning and control, but also for understanding the computational processes performed by other regions of the central nervous system, such as the spinal cord. Although some deviations between the population vector and hand motion have been observed<sup>15</sup>, they have been small and difficult to interpret.

Considerable insight into human motor performance and learning has been gained from studies of planar limb movement where the arm is oriented in the horizontal plane, hand motion is generated only by flexion and extension motions at the shoulder and elbow, and the mechanics of movement can be easily estimated<sup>16–21</sup>. The mechanics of these planar movements are anisotropic with large variations that are dependent on movement direction<sup>17</sup>. We addressed whether the activity of MI neurons at the population level was influenced by these mechanical anisotropies.

We trained monkeys to make planar movements with roughly straight hand trajectories to spatial targets (Fig. 1a). We recorded the activity of neurons in the left contralateral MI of monkeys while they made reaching movements with their right hand from a central target to eight spatial targets that were located on the circumference of a circle. The activity of 214 neurons was found to be unimodally tuned to the direction of movement (62, 22 and 130 in monkeys *a*, *b* and *c*, respectively). As shown previously, cell activity was broadly tuned to the direction of movement. Figure 1b shows the activity of a typical cell in MI during the task where maximal activity occurred when the monkey moved its hand to the right and towards itself (PD = 326°).

We constructed population vectors from our cell sample and compared these vectors to the actual directions of hand motion. Population vectors during movement tended to be biased towards one of two directions: away and left, or towards and right (Fig. 2a). Thirteen of the 16 population vectors did not point in the direction of hand motion (Fig. 2b). There was no significant correlation

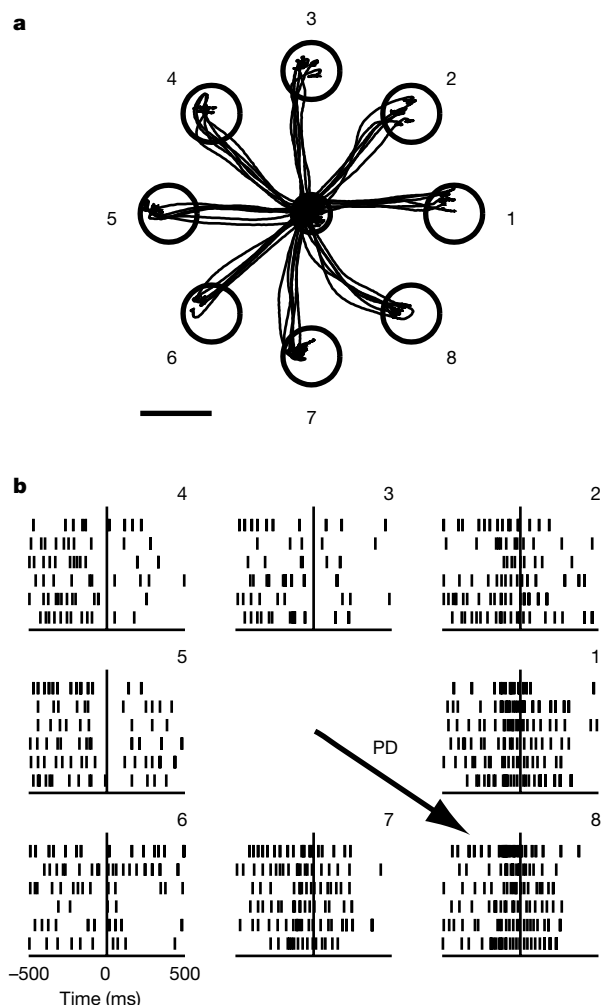
\* Present address: Department of Psychology, The University of Western Ontario, London, Ontario N6A 5C2, Canada.

between deviations in the population vector and deviations in hand movement relative to target direction ( $r^2 = 0.07$ ;  $P > 0.05$ ). Furthermore, there were large fluctuations in population vector length, varying from 56 to 145% of the mean value across all directions. This modulation of the magnitude of the population vector occurred despite the fact that hand movements were of similar magnitude.

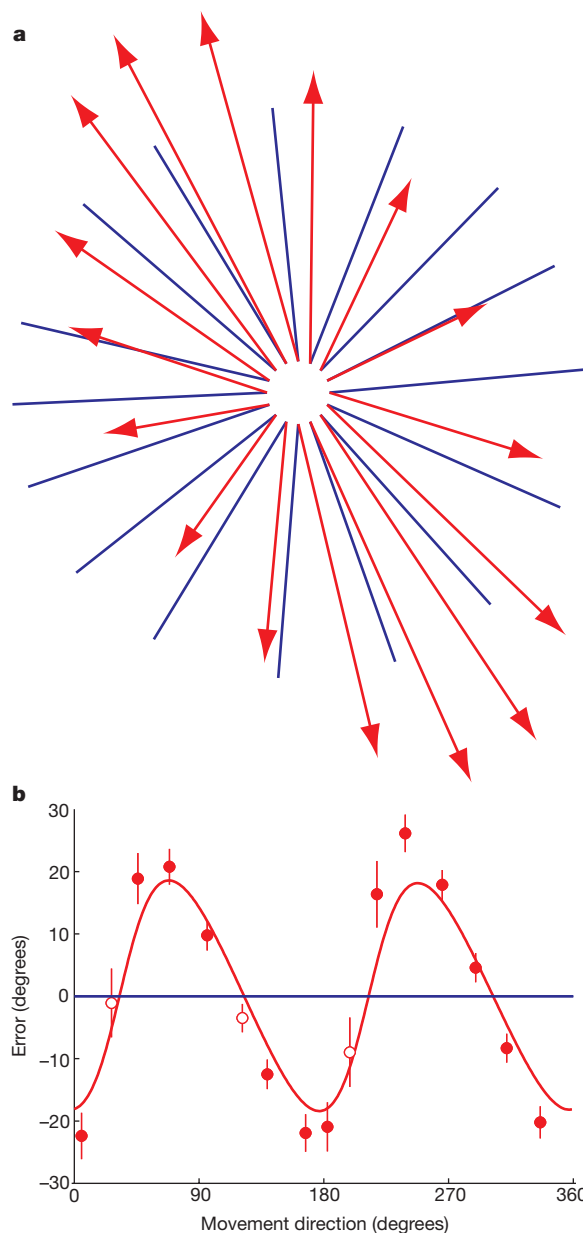
Population vectors will tend to point in the direction of movement if certain conditions are met: no coupling between a cell's PD and the magnitude of its tuning, and a uniform distribution of PDs<sup>7,14</sup>. We examined the former condition by comparing cell modulation against PD. Cell modulation was defined as the highest mean discharge rate before and during movement (reaction time (RT) + movement time (MT)) measured for any movement direction minus its lowest mean discharge rate for any movement direction. Mean cell modulation for a given movement direction was defined from all cells with PDs within  $\pm 22.5^\circ$  of that direction. We found no significant correlation between peak joint power and mean cell modulation associated with each movement direction. ( $r^2 = 0.04$ ;  $P > 0.05$ ; see Supplementary Information).

We examined the distribution of PDs for our sample by segregating cells into 16 groups on the basis of their PD (bin size =  $22.5^\circ$ ) and

then plotting the numbers of cells in each bin against movement direction (Fig. 3a). Spatial directions were not equally represented across the cell sample and the distribution of PDs was found not to be uniform, particularly when compared against a bimodal distribution (bimodal distribution,  $P < 10^{-5}$ , main axis =  $117\text{--}297^\circ$ ; unimodal distribution,  $P < 0.01$ , mean vector is  $26^\circ$ , Rayleigh test). There were two apparent clusters, one for movements away and left, and another for movements towards and right. Cells recorded in each individual monkey showed similar biases in their distribution



**Figure 1** Cell activity in primary motor cortex (MI) during reaching. **a**, Hand trajectory for six repeat trials to eight spatial targets. Scale bar, 3 cm. **b**, Corresponding neural activity for each movement. Each raster plot shows the discharge pattern of the cell for six repeat trials where each small vertical line denotes the occurrence of an action potential. Long vertical bars denote movement onset. The arrow defines the cell's preferred direction (PD), the direction in which it is maximally active (PD =  $326^\circ$ ).



**Figure 2** Comparison between population vectors and directions of hand movement. **a**, The base of each population vector (red arrow) is attached to the corresponding direction of hand movement (blue line). **b**, Difference between the direction of population vector and the direction of hand movement. Filled circles denote significant differences between the population vector and movement direction at the 1% level; open circles denote no significant difference ( $P > 0.01$ ). The vertical line through each circle represents 1 s.d. in the distribution of re-sampled population vectors (see Methods). The thick red line denotes a theoretical prediction of population vector error for all movement directions on the basis of a simulated population of 1,000 cells with a bimodal distribution similar to that observed in the present study where the tuning of each cell was based on a unimodal von Mises function.

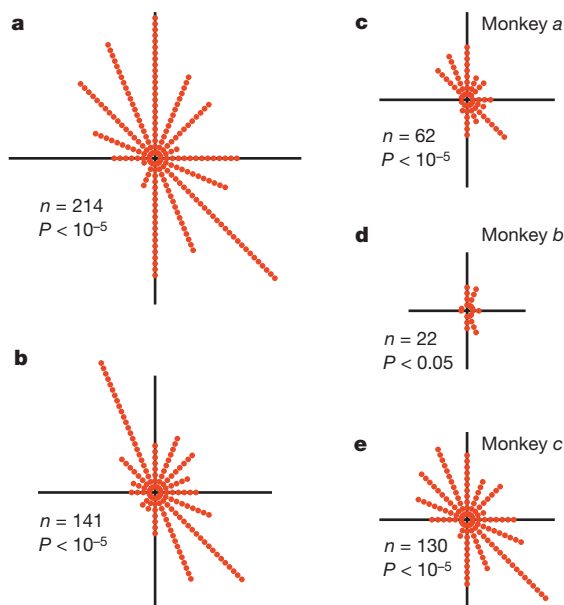
of PDs (Fig. 3c–e). Afferent feedback cannot account for these results as a similar non-uniform distribution was observed when the directional preference of cells was based only on their activity during the reaction time period (Fig. 3b; bimodal distribution,  $P < 10^{-5}$ , main axis =  $118\text{--}298^\circ$ ,  $n = 141$ ; only cells tuned during RT included in this distribution).

Variations in the distribution of PDs were compared to several different kinematic and kinetic features of movement. Hand velocity was essentially the same for movements in the different spatial directions and did not correlate with the distribution of PDs ( $r^2 = 0.10$ ; Fig. 4a). Peak joint velocity (sum of shoulder and elbow joint velocity) varied for movements in different spatial directions and showed a modest correlation with the distribution of PDs ( $r^2 = 0.54$ ; Fig. 4b). Although peak joint torque varied with movement direction, it did not correlate at all with the distribution of PDs ( $r^2 = 0.002$ ; Fig. 4c).

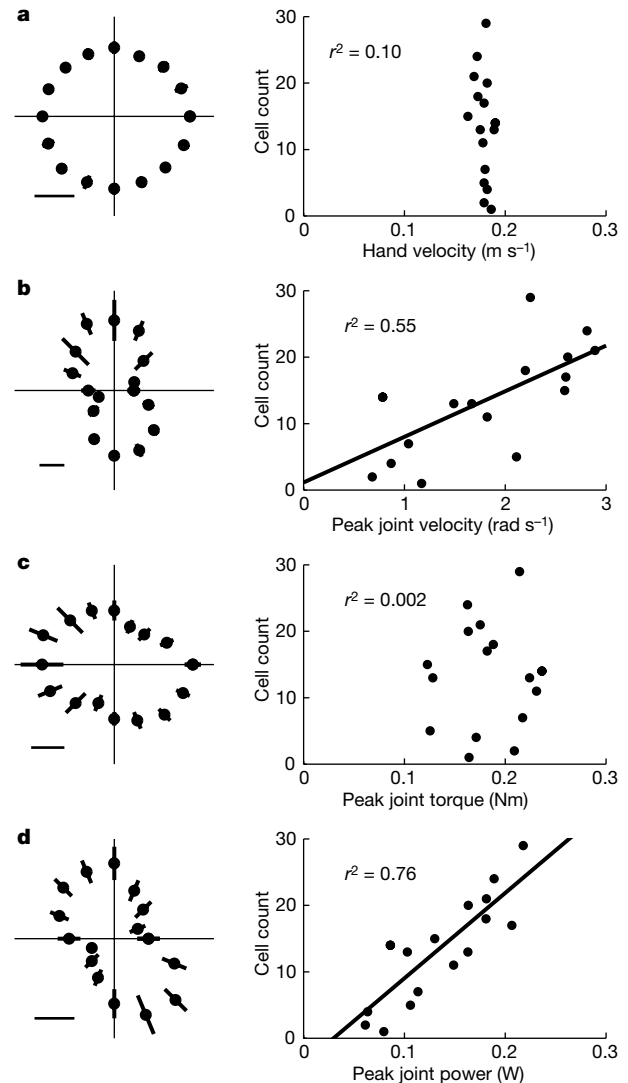
Joint power is a variable that characterizes the mechanical characteristics of the limb as it incorporates two important features of the peripheral motor apparatus: joint torque, which reflects the anisotropic inertial properties of the limb<sup>17</sup>, and joint velocity, which reflects muscle force reduction with shortening velocity<sup>22</sup>. Figure 4d shows that peak joint power varies strongly for movements in different spatial directions. Peak power for movements of the right arm tends to be larger for movements away and to the left (upper left quadrant), and for movements towards and to the right (lower right quadrant). Notably, variations in the distribution of PDs paralleled the anisotropy in peak joint power (Fig. 4d). Movement directions requiring larger peak joint power tended to have more directionally tuned cells than directions requiring less power. A statistically significant correlation ( $r^2 = 0.76$ ) was found between joint power and cell counts associated with each movement direction ( $P < 0.01$ ). Moreover, variations in the length of population vectors for different movement directions correlated strongly with

peak joint power ( $r^2 = 0.94$ ;  $P < 0.001$ ).

The distribution of PDs is somewhat non-uniform when reaching movements are performed with the shoulder near a neutral abduction/adduction angle, but becomes more skewed when reaching with the shoulder abducted at an angle of about  $80^\circ$  (ref. 11). As compared with ref. 11, we found an even stronger non-uniformity when the shoulder was abducted  $90^\circ$  and the arm was maintained in the horizontal plane during reaching. The smaller non-uniformity found in this previous study was probably a result of the arm not being entirely restricted to the horizontal plane and that the



**Figure 3** Directional preference of cells in MI. **a**, Distribution of PDs for the 214 directionally tuned cells during reaction time + movement time (RT + MT) recorded in contralateral motor cortex during movements of the right hand. Each circle represents the directional preference of an individual cell and cells are grouped into 16 equal-sized bins (bin size =  $22.5^\circ$ ). The  $P$ -value in the panel reflects the statistical level at which the distribution was found to be non-uniform as compared with a bimodal distribution. **b**, Distribution of PDs on the basis of the activity of neurons during the reaction time period (300 ms before movement onset). Only neurons directionally tuned before movement onset were included in this sample. **c–e**, Distribution of PDs for neurons recorded in each individual monkey during RT + MT.



**Figure 4** Comparisons between distribution of PDs and various kinematic and kinetic features of movement. Kinesiological values based on all movements in which neural activity was recorded in this task. **a**, Left, polar plot of peak hand velocity relative to direction of hand movement for 16 targets uniformly distributed in space. Mean and s.d. are defined with filled circles and thick lines, respectively. Scale bar,  $0.1 \text{ m s}^{-1}$ . Right, peak hand velocity for a given movement direction plotted against the number of neurons with PDs within  $\pm 11.25^\circ$  of each movement direction. **b**, Variations in peak joint velocity (sum of shoulder and elbow) for movements in different spatial directions. Scale bar,  $1 \text{ rad s}^{-1}$ . Right panel illustrates a significant correlation between joint velocity and cells associated with that direction of movement ( $P < 0.01$ ; black line denotes regression line). **c**, Left, variations in peak joint torque for movements in different directions. Right, lack of a correlation between joint torque and cell counts for each movement direction. Scale bar,  $0.1 \text{ Nm}$ . **d**, Polar plot of peak joint power relative to direction of hand movement. Scale bar,  $0.1 \text{ W}$ . Right, significant correlation between joint power and the distribution of PDs ( $P < 0.01$ ). Variations in peak joint power can explain 76% of the variability in the distribution of PDs.



monkeys moved a pendulum, which requires larger shoulder torques in the frontal as compared with the sagittal plane. Although it is difficult to identify exactly why there is a difference in the distribution of PDs between studies, the non-uniformity observed with the shoulder abducted 90° probably reflects several interacting factors. First, limb mechanics are highly anisotropic. Second, hand movements are generated only by flexion and extension motions at the shoulder and elbow rather than including other degrees of freedom at the shoulder. Third, our previous study, where muscular torque is systematically manipulated during posture, shows that proximal forelimb muscles and neuronal activity in MI tend to be preferentially related to flexor torque at one joint combined with extensor torque at the other<sup>23</sup>. This multi-joint coupling would tend to generate a greater non-uniformity in the distribution of PDs during this reaching task than predicted if neurons were randomly associated to muscle groups at each joint.

A central advance in our study was that the distribution of PDs was correlated to mechanical anisotropies. This correlation suggests that the role of MI in compensating for the mechanical properties of the limb overrides any requirement to maintain a uniform distribution of PDs related to the direction of hand movement. As a result we found that a vector constructed from the ensemble activity of neurons in MI, as described by the population vector hypothesis, does not have to point in the direction of movement for non-human primates to successfully move their hands to spatial targets. Of course, hand motion may be computed from this population of cells either by selecting a sub-population of cells from our sample that are uniformly distributed, or by using more complex computational strategies rather than the population vector method<sup>24,25</sup>.

The success of joint power to predict the distribution of PDs may be partially due to the fact that it is a first approximation of muscle activation during reaching. Muscle activation reflects not only force (or torque) but also velocity as force production drops markedly with muscle shortening velocity<sup>22</sup>. Maximal activities of shoulder and elbow muscles were similar to that observed for MI cells (data not shown). We cannot easily separate these two variables in the present task, and the advantage of using joint power is that it can be easily computed from joint kinematics and inverse dynamics. It is important to note that the present results should not be construed as evidence that neural activity in MI explicitly codes joint power (or even muscle activation) because many variables influence cell activity<sup>5,10–13</sup>. However, this study does emphasize that the complex role played by MI in controlling multi-joint movements cannot be captured easily with a hand-based framework; further progress in this field will require peripheral-based paradigms that consider the mechanics of the limb. □

## Methods

### Experimental paradigm

Three male rhesus monkeys (6–7 kg) were trained to wear a mechanical exoskeleton that permitted flexion and extension motion of the shoulder and elbow joints in the horizontal plane<sup>26</sup>. Monkeys were trained to make movements from a central target (0.8 cm radius) to 16 peripheral targets (1.2 cm radius) uniformly distributed on the circumference of a circle (6 cm radius). A successful trial required the monkey to move between the spatial targets in 220–350 ms (shorter than overall movement time; see below). This movement criteria along with overtraining from months of practice probably resulted in constant movement times for all spatial directions (Fig. 4a), which contrasts with previous observations that hand velocity varies with movement direction in naive human subjects<sup>27</sup>. The monkey was fully trained to perform these movements consistently before neural recordings were initiated. All procedures followed university and national guidelines for animal care.

Recording chambers were implanted surgically under inhalation anaesthetic and we used conventional techniques for extracellular recording of single-neuron activity related to the proximal arm in MI (ref. 11). The territory where cell activity was sampled in MI was similar to the region examined in previous studies on reaching<sup>11</sup>. Neurons recorded in the task were located in the rostral bank and crown of the central sulcus where trains of electrical stimulation (11 pulses, 333 Hz, 0.2 ms pulse width, range 5–50  $\mu$ A) elicited movement of the shoulder or elbow. Cells were examined in the task if they were active during reaching and if they responded predominantly to passive movement of the shoulder and/or elbow. Cells that did not respond to passive movement of any of the forelimb joints were recorded in the task if neighbouring neurons responded only to

passive movement of the shoulder and/or elbow. Neuronal activity was collected along with a number of variables related to hand and joint motion using a custom-designed, data-acquisition system<sup>26</sup>. Cell activity was usually recorded in six repeat trials to 8 of the 16 targets that were presented in a pseudo-random block design. In the first two monkeys (*a* and *b*), neural activity was recorded for movements to 8 targets selected specifically to sample joint-based frameworks (45, 67.5, 90, 180, 247.5, 270, 305 and 327.5°, where 0° is to the right and positive rotation is anticlockwise). Most of the neurons recorded in this study are from monkey *c*, where neural activity was recorded for movements to 8 targets that were distributed symmetrically in space (0, 45, 90, 135, 180, 225, 270 and 315°).

### Data analyses

Mean movement time was 576 ms across all directions and movement onset was defined as 10% of peak hand velocity. Cell activity was based on the number of action potentials recorded 300 ms before movement onset to 575 ms after movement onset (RT + MT). Shoulder and elbow muscular torques were calculated using inverse dynamics<sup>26,28</sup>. Joint power was computed by multiplying each joint's muscular torque by its angular velocity<sup>29</sup> and summing these values for the shoulder and elbow together at 5 ms time intervals. Peak values for each variable equalled the largest summed value for each movement direction.

Most of the cells in this study were recorded from monkey *c* where movement directions were equally distributed in space. The directional preferences of these cells were estimated using standard trigonometric techniques where the directional bias of the cell was defined by a mean vector whose orientation defines the cell's preferred movement direction<sup>11,30</sup>. For neural data recorded in monkeys *a* and *b*, the spatial targets were distributed non-symmetrically. We adapted techniques to describe the area of planar objects to identify the cell's mean vector for each block of eight trials (see Supplementary Information). A standard bootstrap technique was used to define whether the directional tuning of each cell, as defined by the mean vector, was statistically significant at the 1% level<sup>11</sup>.

We tested whether the distribution of PDs was uniform as compared to either a unimodal or bimodal distribution. In each case, a mean vector was computed from the observed distribution of PDs and the statistical test identified whether a random sample drawn from a uniform distribution (all directions equally represented) could generate a vector of similar or greater length (statistics based on 1,000,000 repeat samples).

Population vectors were computed for the cell population using standard techniques<sup>3</sup>. As neural activity was recorded in only 8 of the 16 movement directions, von Mises tuning functions were used to characterize the activity of a cell for each movement direction relative to the mean discharge across directions (see Supplementary Information). We used a bootstrap method to identify whether the direction of each population vector was significantly different from the corresponding direction of hand motion<sup>3,11</sup>. The direction of hand motion for each target was defined by its position at peak hand velocity relative to its initial position before movement onset. For each movement, populations of 214 cells were re-sampled with replacement from the original distribution, and the population vector for each sample was computed to define a sampling distribution of population vectors for each movement direction (total re-sampled distributions = 1,000). The population vector was identified as significantly different from the direction of hand movement if less than 1% of the re-sampled distributions crossed mean hand direction.

Received 19 April; accepted 24 July 2001.

- Georgopoulos, A. P., Caminiti, R., Kalaska, J. F. & Massey, J. T. Spatial coding of movement: a hypothesis concerning the coding of movement directions by motor cortical populations. *Exp. Brain Res. (Suppl. 7)* 327–336 (1983).
- Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal population coding of movement direction. *Science* **233**, 1416–1419 (1986).
- Georgopoulos, A. P., Kettner, R. E. & Schwartz, A. B. Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of arm movement by a neural population. *J. Neurosci.* **8**, 2928–2937 (1988).
- Georgopoulos, A. P. Current issues in directional motor control. *Trends Neurosci.* **18**, 506–510 (1995).
- Caminiti, R., Johnson, P. B. & Urbano, A. Making arm movements in different parts of space: dynamic aspects in the primate motor cortex. *J. Neurosci.* **10**, 2039–2058 (1990).
- Moran, D. W. & Schwartz, A. B. Motor cortical representation of speed and direction during reaching. *J. Neurophysiol.* **82**, 2676–2692 (1999).
- Mussa-Ivaldi, F. A. Do neurons in the motor cortex encode movement direction? An alternative hypothesis. *Neurosci. Lett.* **91**, 106–111 (1988).
- Todorov, E. Direct cortical control of muscle activation in voluntary arm movements: a model. *Nature Neurosci.* **3**, 391–398 (2000).
- Kalaska, J. F., Scott, S. H., Cisek, P. & Sergio, L. E. Cortical control of reaching movements. *Curr. Opin. Neurobiol.* **7**, 849–859 (1997).
- Kalaska, J. F., Cohen, D. A. D., Hyde, M. L. & Prud'homme, M. A comparison of movement direction-related versus load direction-related activity in primate motor cortex, using a two-dimensional reaching task. *J. Neurosci.* **9**, 2080–2102 (1989).
- Scott, S. H. & Kalaska, J. F. Reaching movements with similar hand paths but different arm orientations. I. Activity of individual cells in motor cortex. *J. Neurophysiol.* **77**, 826–852 (1997).
- Shen, L. & Alexander, G. E. Preferential representation of instructed target location versus limb trajectory in dorsal premotor area. *J. Neurophysiol.* **77**, 1195–1212 (1997).
- Kakei, S., Hoffman, D. S. & Strick, P. L. Muscle and movement representations in the primary motor cortex. *Science* **285**, 2136–2139 (1999).
- Sanger, T. D. Theoretical considerations for the analysis of population coding in motor cortex. *Neural Comput.* **6**, 12–21 (1994).
- Scott, S. H. & Kalaska, J. F. Motor cortical activity is altered by changes in arm posture for identical hand trajectories. *J. Neurophysiol.* **73**, 2563–2567 (1995).
- Hollerbach, J. M. & Flash, T. Dynamic interactions between limb segments during planar arm movement. *Biol. Cybern.* **44**, 67–77 (1982).
- Mussa-Ivaldi, F. A., Hogan, N. & Bizzi, E. Neural, mechanical, and geometric factors subserving arm posture in humans. *J. Neurosci.* **5**, 2732–2743 (1985).

18. Karst, G. M. & Hasan, Z. Timing and magnitude of electromyographic activity for two-joint arm movements in different directions. *J. Neurophysiol.* **66**, 1594–1604 (1991).
19. Shadmehr, R. & Mussa-Ivaldi, F. A. Adaptive representation of dynamics during learning of a motor task. *J. Neurosci.* **14**, 3208–3224 (1994).
20. Gomi, H. & Kawato, M. Equilibrium-point control hypothesis examined by measured arm-stiffness during multi-joint movement. *Science* **272**, 117–120 (1996).
21. Sainburg, R. L., Ghez, C. & Kalakanis, D. Intersegmental dynamics are controlled by sequential anticipatory, error correction, and postural mechanisms. *J. Neurophysiol.* **81**, 1040–1056 (1999).
22. Scott, S. H., Brown, I. E. & Loeb, G. E. Mechanics of feline soleus: I. Effect of fascicle length and velocity on force output. *J. Muscle Res. Cell Motil.* **17**, 207–219 (1996).
23. Cabel, D. W., Scott, S. H. & Cisek, P. Neural activity in primary motor cortex related to mechanical loads applied to the shoulder and elbow during a postural task. *J. Neurophysiol.* (in the press).
24. Sanger, T. D. Probability density estimation for the interpretation of neural population codes. *J. Neurophysiol.* **76**, 2790–2793 (1996).
25. Zhang, K., Ginzburg, L., McNaughton, B. L. & Sejnowski, T. J. Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J. Neurophysiol.* **79**, 1017–1044 (1998).
26. Scott, S. H. Apparatus for measuring and perturbing shoulder and elbow joint positions and torques during reaching. *J. Neurosci. Methods* **89**, 119–127 (1999).
27. Gordon, J., Ghilardi, M. F., Cooper, S. E. & Ghez, C. Accuracy of planar reaching movements. II. Systematic extent errors resulting from inertial anisotropy. *Exp. Brain Res.* **99**, 112–130 (1994).
28. Cheng, E. J. & Scott, S. H. Morphometry of *Macaca mulatta* forelimb. I. Shoulder and elbow muscles and segment inertial parameters. *J. Morphol.* **245**, 206–224 (2000).
29. Winter, D. A. Biomechanics of human movement with applications to the study of human locomotion. *Crit. Rev. Biomed. Eng.* **9**, 287–314 (1984).
30. Batschelet, E. *Mathematics in Biology: Circular Statistics in Biology* (Academic, London, 1981).

**Supplementary information** is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

#### Acknowledgements

We thank K. Moore for technical assistance and D. Munoz, M. Pare and K. Rose for helpful comments on the manuscript. S. Chan assisted in the training of one monkey. This research was supported by a CIHR grant and scholarship to S.H.S. and a CIHR fellowship to P.L.G.

Correspondence and requests for materials should be addressed to S.H.S. (e-mail: [steve@biomed.queensu.ca](mailto:steve@biomed.queensu.ca)).

## Rae1 and H60 ligands of the NKG2D receptor stimulate tumour immunity

Andreas Diefenbach, Eric R. Jensen, Amanda M. Jamieson & David H. Raulet

Department of Molecular and Cell Biology and Cancer Research Laboratory, 485 Life Sciences Addition, University of California, Berkeley, California 94720, USA

Natural killer (NK) cells attack many tumour cell lines, and are thought to have a critical role in anti-tumour immunity<sup>1–7</sup>; however, the interaction between NK cells and tumour targets is poorly understood. The stimulatory lectin-like NKG2D receptor<sup>8–13</sup> is expressed by NK cells, activated CD8<sup>+</sup> T cells and by activated macrophages in mice<sup>11</sup>. Several distinct cell-surface ligands that are related to class I major histocompatibility complex molecules have been identified<sup>11–14</sup>, some of which are expressed at high levels by tumour cells but not by normal cells in adults<sup>11,13,15,16</sup>. However, no direct evidence links the expression of these 'induced self' ligands with tumour cell rejection. Here we demonstrate that ectopic expression of the murine NKG2D ligands Rae1 $\beta$  or H60 in several tumour cell lines results in potent rejection of the tumour cells by syngeneic mice. Rejection is mediated by NK cells and/or CD8<sup>+</sup> T cells. The ligand-expressing tumour cells induce potent priming of cytotoxic T cells and sensitization of NK cells *in vivo*. Mice that are exposed to live or irradiated tumour cells expressing Rae1 or H60 are specifically immune to subsequent challenge with tumour cells that lack NKG2D ligands, suggesting application of the ligands in the design of tumour vaccines.

As demonstrated by staining with a tetramerized derivative of the extracellular portion of NKG2D, NKG2D ligands are expressed by most of the tumour cells tested, including various lymphoid, myeloid and carcinoma cell lines (ref. 11; and A. Diefenbach and D. Raulet, unpublished data). Northern blot analysis showed that many of the positive cell lines express Rae1 transcripts, whereas H60 transcripts were limited to only one or two of the cell lines tested (data not shown). Rae1 transcripts have not been detected in normal cells from adult mice<sup>15</sup>, suggesting that these genes are specifically upregulated in tumour cell lines.

To investigate whether tumour cells that express NKG2D ligands stimulate anti-tumour immune responses, we used a retrovirus expression system to ectopically express high levels of Rae1 $\beta$  or H60 in EL4 (a thymoma), RMA (a T-cell lymphoma) and B16-BL6 (a melanoma). These cell lines are all from C57BL/6 (hereafter termed B6) mice and do not normally express NKG2D ligands<sup>11</sup>. Ligand-transduced cells were selected on the basis of staining with NKG2D tetramers. To serve as controls, tumour cells that were transduced with empty retrovirus vector (designated as EL4/–, B16/– and RMA/–) were selected by genomic polymerase chain reaction (PCR) (see Methods).

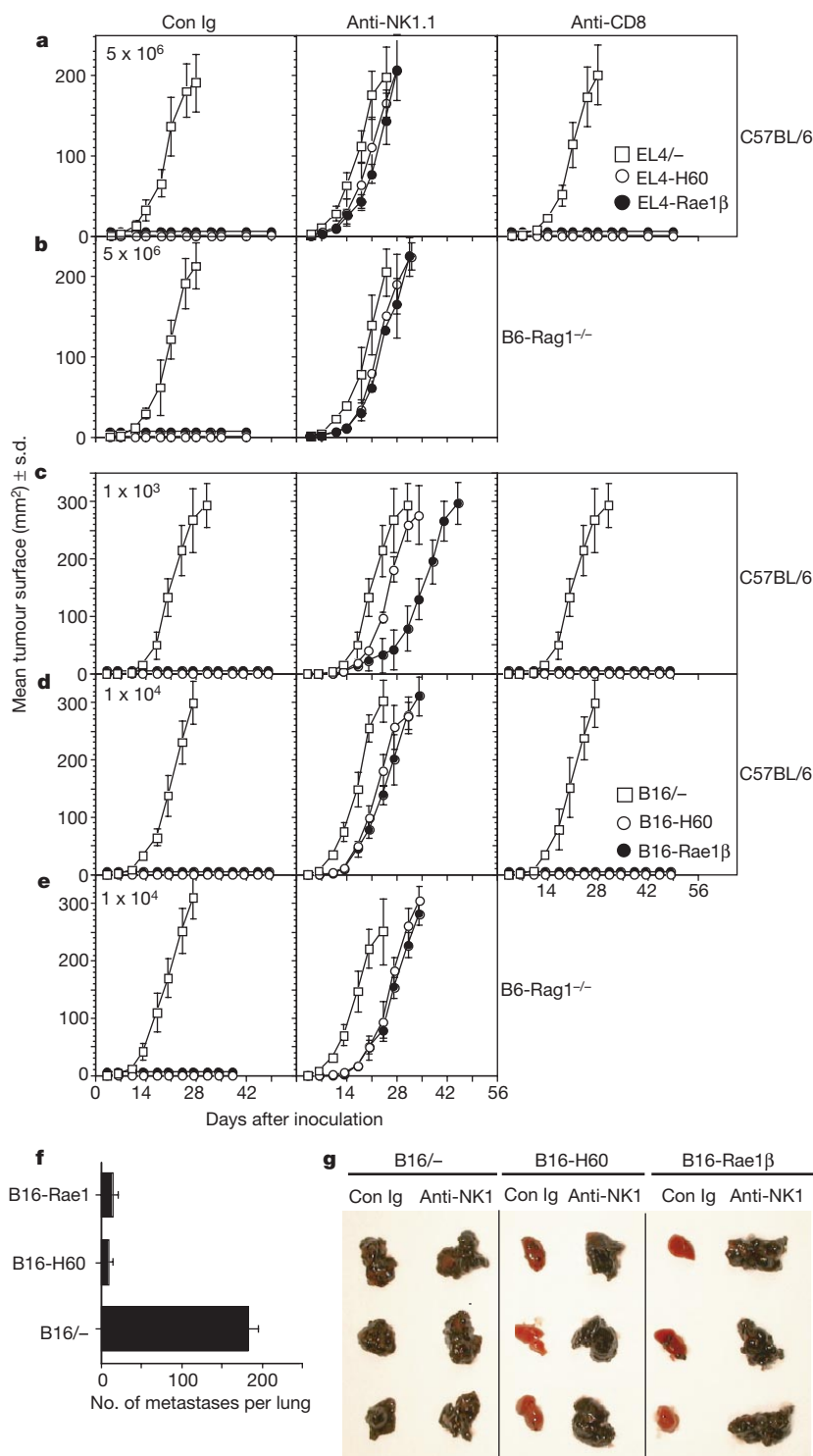
For analysis of the response to EL4 and B16-BL6 tumour cells, groups of five B6 mice were inoculated subcutaneously with syngeneic tumour cell transductants. Control-transduced EL4 or B16-BL6 cells grew progressively at a rate similar to that of untransduced cells (Fig. 1a, c, d, and data not shown), leading to uniform terminal morbidity by about 28 days. Notably, Rae1 $\beta$ - or H60-transduced tumour cells of both types were rejected rapidly and completely, as they failed to yield detectable tumours at any time point (Fig. 1a, c, d). A tenfold increase in the dose of Rae1 $\beta$ - or H60-transduced EL4 cells (to  $5 \times 10^7$  cells) did not change the outcome, whereas a higher dose ( $1 \times 10^5$ ) of Rae1 $\beta$ - or H60-transduced B16-BL6 cells resulted in progressive, although substantially delayed, tumour growth in all the mice, compared with the control-transduced tumour cells (data not shown). Ligand-transduced tumour cells of both types also failed to grow in B6 mice that had been depleted of CD8<sup>+</sup> T cells or in B6-Rag1<sup>–/–</sup> mice—which lack all T and B cells—but grew progressively in normal and B6-Rag1<sup>–/–</sup> hosts that had been depleted of NK1.1<sup>+</sup> cells (Fig. 1a–e). Thus, these doses of Rae1 $\beta$ - or H60-transduced EL4 cells and B16-BL6 cells are rejected rapidly by conventional NK cells without a requirement for T and B cells, including NK1.1<sup>+</sup> T cells or  $\gamma\delta$  T cells. Interestingly, Rae1 $\beta$ - or H60-transduced B16-BL6 cells reproducibly exhibited retarded growth in NK-depleted B6-Rag1<sup>–/–</sup> mice (Fig. 1e). It is possible that a residual response against these cells is mediated by non-lymphoid cells such as macrophages, or by small numbers of NK cells that survive antibody treatment.

Rae1 $\beta$  or H60 expression by B16-BL6 cells reduced the frequency of lung metastases by over tenfold after intravenous injection (Fig. 1f). In another experiment where mice were examined at a later time point, control-transduced B16-BL6 cells formed massive contiguous lung metastases, but ligand-transduced B16-BL6 cells were almost completely rejected (Fig. 1g). NK1.1 depletion before tumour cell inoculation markedly depressed the rejection of the metastases.

Rae1 $\beta$ - or H60-transduced RMA tumour cells were also rejected by B6 mice (Fig. 2). Unlike the responses to the other tumour cells, however, the primary rejection of ligand-transduced RMA cells was mediated by both CD8<sup>+</sup> T cells and NK cells, although the specific outcome depended on the dose of tumour cells. Depletion of both NK1.1<sup>+</sup> cells and CD8<sup>+</sup> T cells was necessary to abrogate rejection of the smallest inoculum of  $10^4$  ligand-transduced tumour cells, whereas depletion of either population allowed tumour growth in at least some animals that were injected with the largest dose ( $10^6$ ) of tumour cells (Fig. 2a). With the intermediate dose of  $10^5$  tumour cells, depletion of CD8 cells allowed tumour cell growth, but NK cell depletion did not (Fig. 2a). Thus, either subset is sufficient for

tumour rejection at the lowest tumour cell dose, CD8 cells (but not NK cells) are sufficient at the intermediate dose, and the two cell types must cooperate to achieve consistent rejection at the highest tumour cell dose. In B6-Rag1<sup>-/-</sup> mice, NK cells were sufficient to reject the Rae1 $\beta$ - or H60-transduced RMA cells at the two lower tumour doses (Fig. 2b), suggesting that NK cells are more active or

effective in these B6-Rag1<sup>-/-</sup> mice than in B6 mice. Parallel analysis of mice that were inoculated with the RMA/S cell line—a major histocompatibility complex (MHC) class I<sup>low</sup> version of RMA cells—confirmed previous reports that these cells are rejected by NK cells and not by CD8<sup>+</sup> T cells<sup>2</sup>, and also demonstrated the efficacy of our NK cell-depletion procedure (Fig. 2a, b).



**Figure 1** EL4 and B16-BL6 tumour cells that express NKG2D ligands are rejected by NK cells in syngeneic mice. Con Ig, control immunoglobulin. **a–e**, C57BL/6 or B6-Rag1<sup>-/-</sup> mice were treated with the indicated antibodies before inoculation with the indicated number of EL4 (**a**, **b**) or B16-BL6 (**c–e**) tumour cell transductants. Tumour growth was

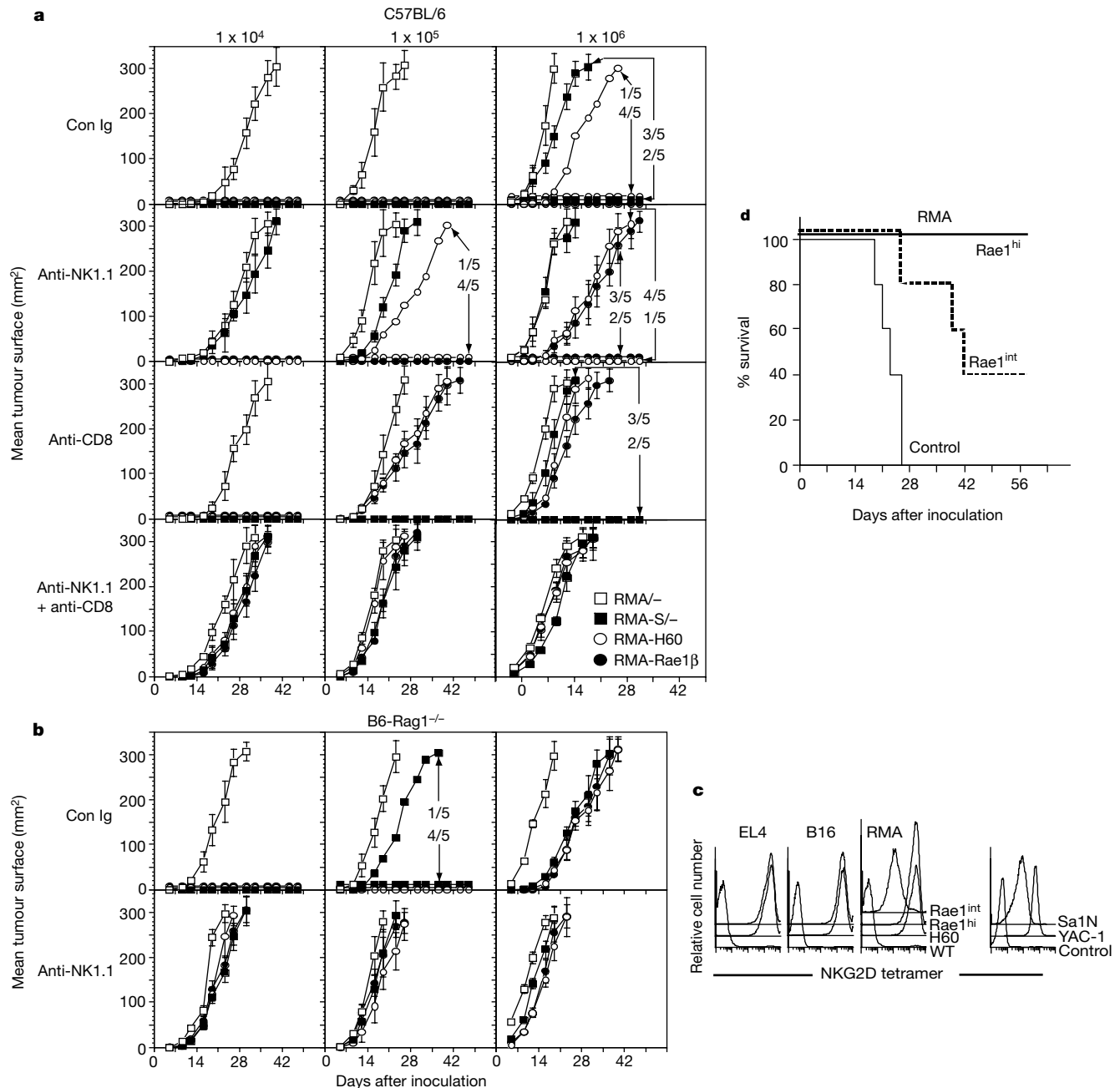
monitored thereafter. Data are representative of at least three experiments. **f**, **g**, B6 mice were injected intravenously with  $3 \times 10^5$  B16-BL6 transductants. Lung metastases were counted two weeks after the tumour challenge (**f**), or the lung lobes from representative mice in a separate experiment were examined three weeks after tumour challenge (**g**).



Why does NKG2D ligand expression not cause the rejection of most of the tumour cell lines that naturally express these ligands<sup>11</sup>? In some cases, the host response to the ligands may be overwhelmed by an excess tumour load. Furthermore, the response may critically depend on the levels of NKG2D ligands expressed by tumour cells, especially if the NK cells are also subject to inhibitory signalling by means of MHC-specific receptors. Notably, the levels of Rae1 $\beta$  and H60 expression in our transductants were substantially higher than the levels of endogenous NKG2D ligands on most of the naturally expressing tumour cell lines that we tested (Fig. 2c). Indeed, a comparison of varying ligand levels showed that RMA

transductants that had intermediate ligand levels, comparable to most tumour cells<sup>11</sup> (Fig. 2c, data not shown), were less efficiently rejected than were transductants with higher ligand levels (Fig. 2d). Therefore, the antitumour response to naturally arising tumour cells probably varies depending on the level of NKG2D ligands that are expressed.

To address whether prior immunization with tumour cells that express NKG2D ligands induces protective immunity to ligand-negative tumour cells, mice that had previously rejected Rae1 $\beta$ - or H60-transduced tumour cells (EL4, B16-BL6 or RMA cells) were re-challenged with corresponding ligand-negative tumour cells 8–12



**Figure 2** RMA cells expressing NKG2D ligands are rejected by NK cells and CD8 T cells in syngeneic mice. **a**, **b**, B6 (**a**) or B6-Rag1<sup>-/-</sup> (**b**) mice were pretreated with the indicated antibodies and inoculated subcutaneously with the indicated numbers of RMA transductants or RMA/S cells. Fractions are indicated for instances where tumours grew in some, but not all, mice. Data are representative of four experiments. Con Ig, control immunoglobulin. **c**, Levels of NKG2D ligands determined by staining with fluorescently

labelled tetrameric NKG2D. Tumour cell transductants, the YAC-1 lymphoma and the Sa1N fibrosarcoma are compared. WT, wild type. **d**, The anti-tumour response depends on Rae1 $\beta$  levels. Survival of B6 mice that had been inoculated with 1  $\times$  10<sup>5</sup> RMA cells, or RMA cells expressing high (Rae1<sup>hi</sup>) or intermediate (Rae1<sup>int</sup>) levels of Rae1 $\beta$ , is shown. Terminally moribund mice were killed. Data are representative of two experiments.

weeks after the first exposure. The ligand-negative tumour cells grew progressively in naive B6 mice, but were rejected by the mice that had been exposed previously to the corresponding Rae1 $\beta$ - or H60-transduced tumour cells (Fig. 3a). The immunity was specific, as mice that had previously rejected ligand-transduced tumour cells of each type were immune only to untransduced tumour cells of the same type (Fig. 3b). Some delay was observed in the rejection of RMA cells by mice that were primed with ligand-transduced EL4 cells (Fig. 3b), and this may be in line with evidence indicating a common origin of EL4 and RMA cells<sup>17,18</sup>. However, the cross-protection was only partial, and was absent or very weak in the reciprocal situation, suggesting that our cell lines have an independent origin, or that most of the response is focused on antigens that are unique to each cell line. We conclude that ligand-expressing tumour cells specifically vaccinated the mice against ligand-negative tumour cells of the same type.

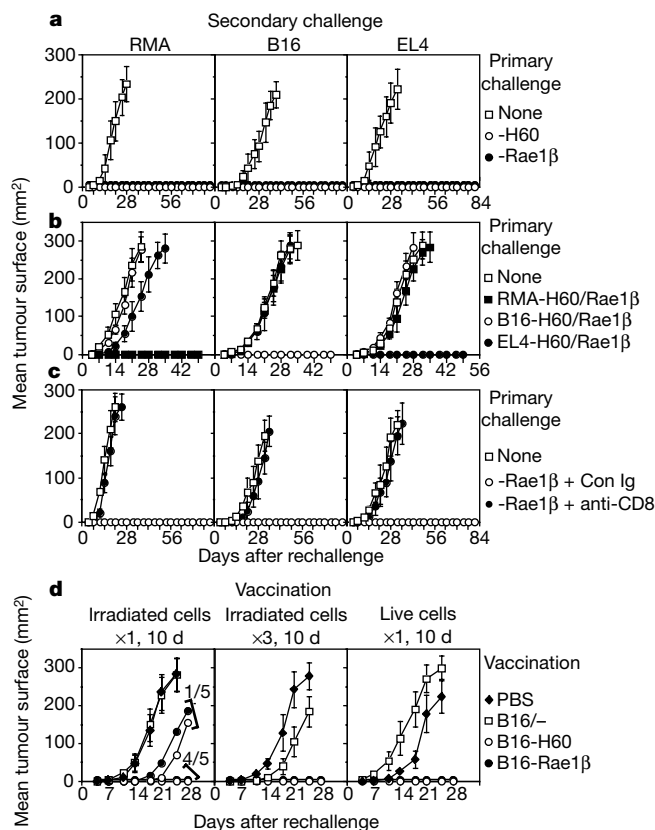
Primary rejection of ligand-transduced EL4 and B16-BL6 cells by naive mice was mediated by NK cells (Fig. 1), but it was not expected that NK cells could provide a specific 'memory' immune

response. Indeed, no immunity to any of the tumour cells developed in Rag1<sup>-/-</sup> mice (Supplementary Information Fig. 1). Furthermore, normal mice that were pretreated with anti-CD8 antibody before the initial exposure to all three types of ligand-transduced tumour cells were unable to reject ligand-negative tumour cells on re-challenge (Fig. 3c), despite the fact that the ligand-transduced tumour cells had been rejected in each case (Figs 1 and 2). Pretreatment of mice with control immunoglobulin did not alter the memory response to ligand-negative tumour cells. Thus, although CD8<sup>+</sup> T cells were not required to reject the primary inoculum of ligand-transduced EL4 and B16-BL6 cells, they were essential for a protective immune response against untransduced tumour cells.

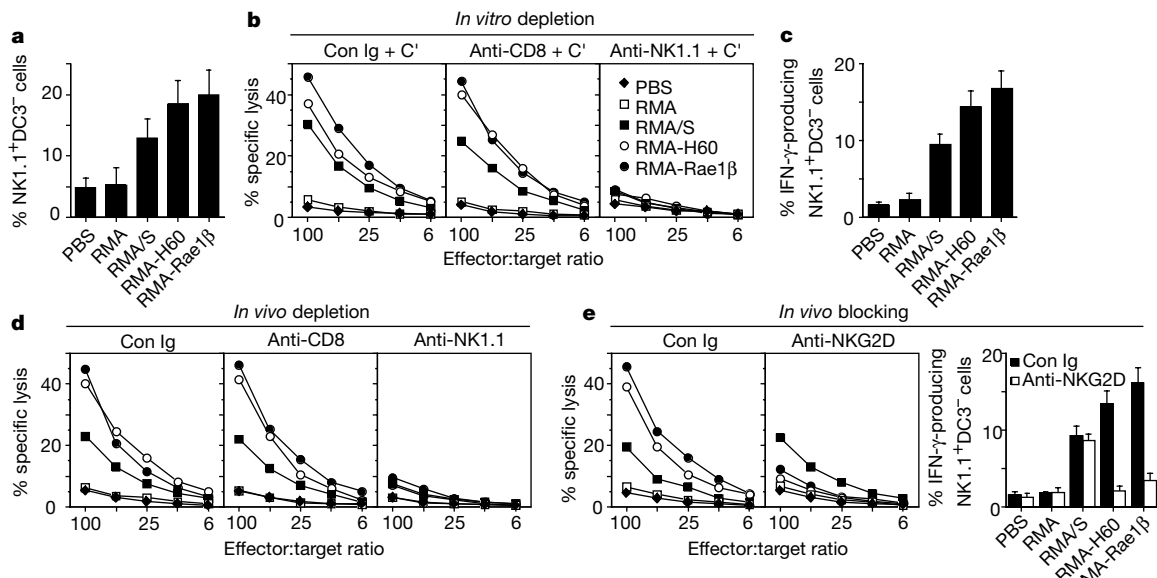
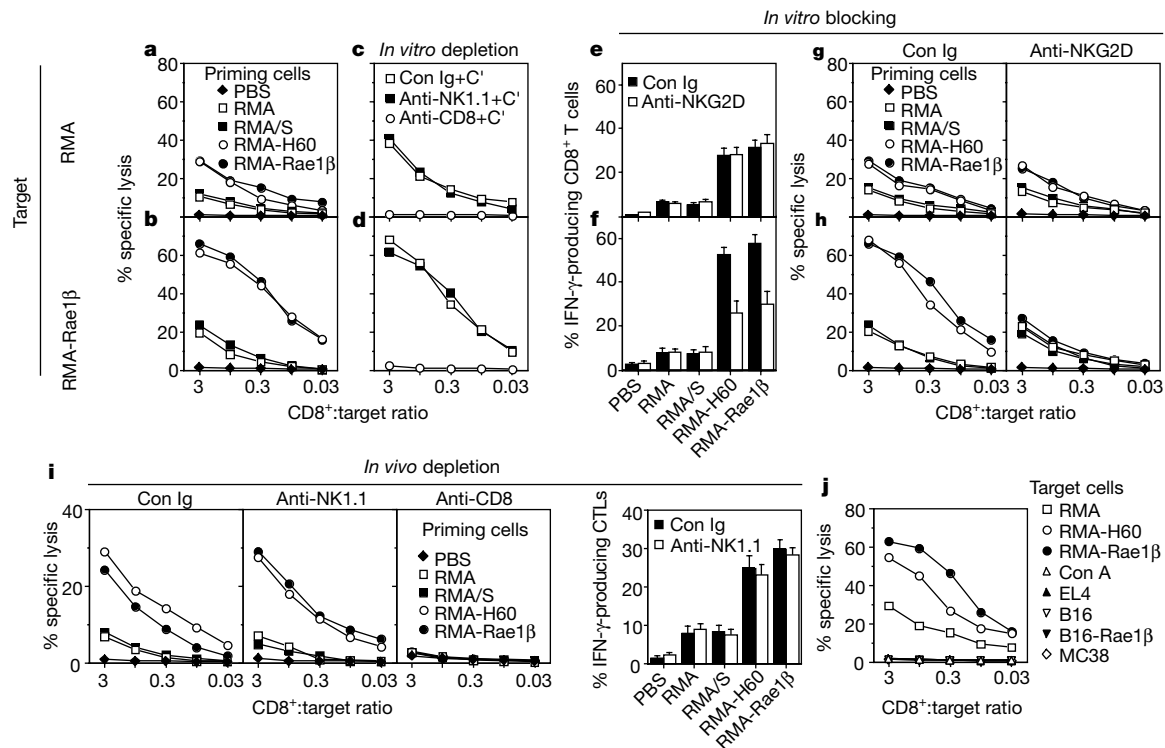
It was possible that immunity arose in this system because NK cell rejection of the transduced tumour cells resulted in excess tumour cell debris, thereby enhancing 'crosspriming' of tumour-specific CD8<sup>+</sup> T cells<sup>19</sup>. Indeed, irradiated untransduced RMA cells, which are expected to die *in vivo* and yield considerable cell debris, are reportedly effective in vaccinating mice<sup>20</sup>. However, this was not the case with irradiated ligand-negative B16-BL6 tumour cells<sup>21</sup>, as no immunity resulted when mice were vaccinated once or three times in succession with irradiated control-transduced B16-BL6 tumour cells (Fig. 3d). Mice that were vaccinated in parallel with irradiated ligand-transduced B16-BL6 cells did develop immunity to untransduced tumour cells, suggesting that ligand expression results in priming of tumour-specific T cells independent of the generation of cell debris by NK cells. Consistent with this conclusion, irradiated ligand-transduced B16-BL6 cells were effective at priming a tumour-specific cytotoxic T-lymphocyte (CTL) response, even in mice that had been previously depleted of NK cells, whereas control-transduced B16-BL6 cells failed to prime CD8<sup>+</sup> T-cell responses (see below and Supplementary Information Fig. 2). Similar results were obtained in the RMA system (Fig. 4).

Consistent with the role of CD8<sup>+</sup> T cells in responses to ligand-transduced tumour cells, irradiated Rae1 $\beta$ - or H60-transduced tumour cells were more effective than ligand-negative tumour cells in priming tumour-specific CTLs *in vivo*, as tested by restimulating with ligand-transduced cells and testing effector function *in vitro*. Priming with irradiated ligand-transduced RMA cells, for example, increased RMA target cell lysis substantially, and augmented the percentage of RMA-specific interferon (IFN)- $\gamma$  producing CD8<sup>+</sup> T cells severalfold (Fig. 4a, e). The effector cells were considerably more active against ligand-transduced RMA cells (Fig. 4b, f), and this enhancement could be completely blocked with anti-NKG2D antiserum (compare Fig. 4e, g with f, h), indicating that the NKG2D ligand interaction enhances not only CTL induction *in vivo*, but also the effector activity of the CTLs. The enhanced induction of CTLs, which were conventional CD8<sup>+</sup>NK1.1<sup>-</sup> cells (Fig. 4c–f), occurred even in mice that had been previously depleted of NK1.1<sup>+</sup> cells (Fig. 4i), demonstrating that ligand-dependent augmentation of the CTL response does not depend directly or indirectly on NK cell activity.

Importantly, the CTLs were tumour-cell specific, as they did not lyse any of three other tumour cells (EL4, B16-BL6 or MC38), nor did they lyse Rae1 $\beta$ -transduced B16-BL6 cells (Fig. 4j). Furthermore, the CTLs did not lyse syngeneic concanavalin A-induced T-cell blasts, suggesting that priming with ligand-transduced tumour cells does not break self-tolerance of CTLs (Fig. 4j). Highly comparable results were obtained in the B16-BL6 tumour cell system, although the responses were generally weaker, probably because of the low expression of class I MHC molecules by these cells (Supplementary Information Fig. 2). Similar results were also obtained with CTLs derived from mice that had previously rejected live ligand-transduced tumour cells (data not shown). Thus, tumour cell expression of NKG2D ligands results in a significant enhancement in the priming of tumour-specific CTLs *in vivo*, as well as in the activation of preformed CTLs. These data are in line



**Figure 3** Vaccination with ligand-expressing tumour cells confers specific immunity to the corresponding ligand-negative tumour cells. **a**, B6 mice that had previously rejected ligand-transduced tumour cells ( $5 \times 10^6$  EL4,  $1 \times 10^4$  B16-BL6 or  $1 \times 10^4$  RMA transductants) were inoculated subcutaneously with control-transduced tumour cells of the same type (EL4,  $5 \times 10^6$ ; B16-BL6,  $1 \times 10^4$ ; RMA,  $1 \times 10^5$ ). Primary exposure occurred 8–12 weeks before challenge. **b**, B6 mice that had been vaccinated 12 weeks earlier with each ligand-transduced tumour cell type were injected with the same or different ligand-negative cell lines. The primary and secondary tumour doses were as in **a**. **c**, Depletion of CD8<sup>+</sup> cells during the primary challenge prevents development of immunity. Re-challenge with control-transduced tumour cells occurred 8–12 weeks after vaccination. Naive B6 mice were challenged in parallel. **d**, B6 mice were injected with PBS or were vaccinated with  $5 \times 10^6$  irradiated (once or three times) or  $1 \times 10^4$  living, transduced or untransduced B16-BL6 tumour cells. Mice were challenged 10 days later with  $10^4$  untransduced B16-BL6 cells in the opposite flank. Data are representative of two experiments.



**Figure 5** RMA cells expressing NKG2D ligands stimulate NK cell recruitment and activation *in vivo*. Groups of five B6 mice were injected intraperitoneally with  $5 \times 10^6$  irradiated RMA transductants, RMA/S cells or PBS. Peritoneal wash cells were recovered two days later. Con Ig, control immunoglobulin. Compared with RMA cells, ligand-transduced RMA cells stimulated elevated percentages of NK (NK1.1 $^+$ CD3 $^+$ ) cells (**a**), which exhibited enhanced capacity to lyse YAC-1 target cells (**b**), and produced IFN- $\gamma$  when stimulated with YAC-1 target cells (**c**). Effector cells were destroyed by complement

(C') lysis with anti-NK1.1 but not anti-CD8 antibody (**d**), and pretreatment of mice with anti-NK1.1 antibody but not with anti-CD8 antibody prevented induction of cytotoxic activity (**e**). **e**, NK cell induction by ligand-transduced cells was blocked by injection of a non-depleting anti-NKG2D antiserum, but not by a control antiserum, just before tumour cell inoculation. The response to RMA/S cells was unaffected. Lysis of YAC-1 tumour cells and production of IFN- $\gamma$  after stimulation with YAC-1 target cells was assayed. Data are representative of two experiments.



with evidence that NKG2D engagement co-stimulates human CTL responses *in vitro*<sup>22</sup>, and extend these findings to *in vivo* responses to tumour cells.

Induction of NK cell activity *in vivo* by tumour cells that express Rae1 $\beta$  or H60 was also demonstrable. When inoculated intraperitoneally in naive mice, irradiated Rae1 $\beta$ - or H60-transduced RMA tumour cells increased the number of peritoneal NK cells by 2–4-fold within two days (Fig. 5a), and substantially enhanced the cytotoxic activity of the cells against YAC-1 target cells (Fig. 5b, left panel), whereas control-transduced RMA cells had no effect. A substantially higher fraction of these NK cells also produced IFN- $\gamma$  after *in vitro* stimulation with YAC-1 tumour cells (Fig. 5c). As previously reported, class I-deficient RMA/S cells also induced the local accumulation of NK cells and enhanced their functional activity<sup>23</sup> (Fig. 5a–c). In either case, the cytotoxic activity was nearly abolished by pretreatment of effector cells with anti-NK1.1 plus complement or by *in vivo* depletion of NK1.1<sup>+</sup> cells before tumour cell inoculation (Fig. 5b, d). In contrast, depletion of CD8<sup>+</sup> cells before tumour cell inoculation or immediately before the cytotoxicity assay had no effect (Fig. 5b, d), and similar cytotoxic activity was induced in Rag1<sup>−/−</sup> and wild-type mice (Supplementary Information Fig. 3). Thus, the effector cells are conventional NK cells and do not depend on CD8<sup>+</sup> cells for their induction. Similar cytotoxicity and cytokine data were obtained with ligand-expressing B16-BL6 and EL4 cells (Supplementary Information Fig. 4). Therefore, expression of NKG2D ligands, like MHC deficiency, provokes NK cell recruitment and sensitization *in vivo*.

A non-depleting anti-NKG2D antiserum injected *in vivo* almost completely blocked the ligand-dependent induction of NK cytotoxicity and cytokine production against YAC-1 cells (Fig. 5e). The effect of the antibody was specific, as NK cell induction by class I-deficient RMA/S cells, which do not express NKG2D ligands, was unaffected by the anti-NKG2D antibody. Therefore, the induction process with ligand-transduced cells required interactions with the NKG2D receptor *in vivo*.

Our results demonstrate that NKG2D ligand expression and consequent activation of NK cells and CD8<sup>+</sup> T cells can impose a substantial barrier to the establishment of tumours *in vivo*. The ligands activate NK cells and T cells through NKG2D, which associates in the membrane with KAP/DAP10—an adapter signalling protein that is thought to deliver co-stimulatory signals<sup>10</sup>. It is unknown whether engagement of NKG2D in NK cells results in direct stimulation or supplies a co-stimulatory signal that acts in conjunction with signals from other stimulatory receptors<sup>24,25</sup>. Nevertheless, NKG2D receptor engagement clearly enhances the effective NK cell response against the three tumour cell lines that we tested. Furthermore, ligand expression by tumour cells also strongly enhanced the generation of tumour-specific CD8 T cells. As this enhancement also occurred in mice that were devoid of NK cells (Figs 2a and 4), and considering that activated CD8<sup>+</sup> T cells express NKG2D, we propose that CD8<sup>+</sup> T-cell activation is enhanced as a consequence of direct interactions with ligand-expressing tumour cells. It remains possible that the priming of CD8<sup>+</sup> T cells is accomplished indirectly as a result of ligand-expressing tumour cells interacting with NKG2D-expressing macrophages<sup>11</sup>. In either case, the results suggest that NKG2D ligands provide an example of innate immune stimuli that function to enhance the adaptive immune response<sup>26</sup>. The effectiveness of ligand transduction of tumour cells in stimulating an anti-tumour response and protective immunity to tumour re-challenge suggests that ligand-expressing cells may have applications in tumour therapy and the development of tumour vaccines. The strong response against B16-BL6 cells is particularly notable in this regard, given that the BL6 variant was selected for high invasiveness and is poorly immunogenic<sup>27</sup>. Indeed, other manipulations such as ectopic B7 expression or CTLA4 blockade do not, by themselves, result in rejection of B16-BL6 cells<sup>28</sup>. The finding that typical levels of NKG2D ligands naturally

found on most tumour cell lines are suboptimal in inducing anti-tumour immunity (Fig. 2c, d) raises the possibility that immunity to such tumours can be boosted by engineering cells with higher levels of ligands. The effectiveness of this approach against pre-established tumours and naturally arising tumours remains to be established. □

## Methods

### Ectopic expression of NKG2D ligands

Three NKG2D ligand-negative cell lines (EL4, a B6 thymoma; RMA, a B6 T lymphoma derived from the Rauscher virus-induced RBL-5 cell line<sup>27</sup>; and B16-BL6, a B6 melanoma derived from the B16-F0 cell line<sup>27</sup>) were retrovirally transduced as described<sup>11</sup>. The retroviral vectors containing the H60 or Rae1 $\beta$  complementary DNAs used for these experiments did not direct synthesis of green fluorescent protein (GFP) or any other selection marker. Transduced cells expressing equivalent high levels of the NKG2D ligands were sorted after staining with a tetrameric soluble version of NKG2D<sup>11</sup>. Control staining was performed with an irrelevant tetramer of the T22 class Ib molecule. Control tumour cells were infected with empty retrovirus, and transduced clones were identified by PCR with primers corresponding to the murine stem cell virus (MSCV) 5' and 3' long terminal repeat (5' primer: GTCCTCCGATAGACTGCGTCGCCCCGGG; 3' primer: GCTTGCCAAACCTACAGGTGGGG). Approximately 100–150 clones with integrated provirus were pooled and used as control tumour cells (designated as EL4/−, B16-BL6/− and RMA/−).

### Antibody depletion, tumour inoculation and re-challenge

We used five mice per group for all tumour rejection experiments. C57BL/6J (B6) and B6-Rag1<sup>−/−</sup> mice were purchased from Jackson Laboratories, and the B6-Rag1<sup>−/−</sup> mice were bred in our animal facilities under specific pathogen-free conditions. All mice were used between 8 and 18 weeks of age. NK cells and CD8<sup>+</sup> T cells were depleted<sup>29</sup> by intraperitoneal injection of 200  $\mu$ g monoclonal antibody (PK136 against NK1.1 and 2.43 against CD8) at day −1, 1, 8, 15 and 22. Control mice received the equivalent amounts of normal mouse immunoglobulin-G. Depletions were confirmed in lymph node and spleen cells 3 weeks after tumour challenge by flow cytometry using non-competing antibodies. In general, less than 1.5% of the depleted cell population could be detected in spleen and lymph nodes. Tumour cells were injected subcutaneously in 100  $\mu$ l PBS in the right flank. We monitored tumour development by measuring the tumour size twice weekly with a metric caliper. For the metastasis assay, 3  $\times$  10<sup>5</sup> B16-BL6 cells were injected intravenously through the tail vein in groups of five mice, and lung metastases were examined 14–21 days later. For the re-challenge experiments, groups of five mice that had completely rejected the initial tumour (8–12 weeks after initial tumour challenge) were injected in the opposite flank with the respective untransduced or control-transduced tumour cells (that is, lacking NKG2D ligands).

### Vaccination and *ex vivo* analysis of CTL activation

Transduced B16-BL6 tumour cells were irradiated (160 Gray (gy)) and 5  $\times$  10<sup>6</sup> cells were injected, in groups of five mice, in 100  $\mu$ l PBS subcutaneously in the left flank either once at day −10 or three times at day −10, −7 and −4 as indicated. Some mice received injections of unirradiated tumour cells at day −10. At day 0, mice were challenged with 10<sup>4</sup> untransduced B16-BL6 cells in 100  $\mu$ l PBS in the opposite flank. For the *ex vivo* analysis of CTL activity, groups of five mice received two vaccinations (day 0 and 3) with irradiated (160 gy) RMA or B16-BL6 tumour cell transductants. Two to three weeks after vaccination, mice were killed and pooled splenocytes were restimulated for 5 days with the respective Rae1 $\beta$ -transduced tumour cells. The cells were expanded for another three days before testing cytotoxicity and IFN- $\gamma$  production. In some experiments NK1.1<sup>+</sup> or CD8<sup>+</sup> T cells were depleted before the cytotoxicity assay by complement (C')-mediated lysis as described<sup>29</sup>. *In vitro* blocking of NKG2D was carried out as described previously, using a rat anti-NKG2D antiserum and a control rat serum<sup>11</sup>.

### *Ex vivo* analysis of NK cell activation

Tumour cells were irradiated (120 gy) and injected intraperitoneally as described<sup>23</sup>. We collected peritoneal cells after 72 h (ref. 23). The percentages ( $\pm$  s.d.) of NK cells in the peritoneal cavity of individual mice were quantified by flow cytometry and electronic gating on lymphocytes (low forward and side light scatter). Peritoneal wash cells were pooled for cytotoxicity and IFN- $\gamma$  assays. In some experiments NK1.1<sup>+</sup> or CD8<sup>+</sup> T cells were depleted before the cytotoxicity assay by complement-mediated lysis as described<sup>29</sup>. For the *in vivo* blocking experiments 100  $\mu$ l of a non-depleting antiserum specific to NKG2D or a control antiserum<sup>11</sup> was injected intraperitoneally. As determined by flow cytometry, the antibody did not deplete NK1.1<sup>+</sup> cells from the peritoneal cavity (data not shown).

### Assays

Cytotoxicity was tested in a standard 4-h <sup>51</sup>Cr release assay<sup>11</sup>. Data are given as the mean of triplicate measurements. Standard deviations were generally less than 5% and are omitted from the figures for reasons of clarity. IFN- $\gamma$  production was assayed by stimulating CTLs or NK cells with an equal number of target cells for 5–7 h and staining intracellular IFN- $\gamma$  as described<sup>30</sup>, with gating on CD8<sup>+</sup>CD3<sup>+</sup> cells or NK1.1<sup>+</sup>CD3<sup>+</sup> cells. Data are presented as the mean ( $\pm$  s.d.) of triplicate measurements.

Received 27 April; accepted 3 August 2001.

- Trinchieri, G. Biology of natural killer cells. *Adv. Immunol.* **47**, 187–376 (1989).
- Karre, K., Ljunggren, H. G., Piontek, G. & Kiessling, R. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature* **319**, 675–678 (1986).
- Seaman, W., Slesinger, M., Eriksson, E. & Koo, G. Depletion of natural killer cells in mice by monoclonal antibody to NK-1.1. Reduction in host defense against malignancy without loss of cellular or humoral immunity. *J. Immunol.* **138**, 4539–4544 (1987).
- van den Broek, M. *et al.* Decreased tumor surveillance in perforin-deficient mice. *J. Exp. Med.* **184**, 1781–1790 (1996).
- Kaplan, D. H. *et al.* Demonstration of an interferon gamma-dependent tumor surveillance system in immunocompetent mice. *Proc. Natl Acad. Sci. USA* **95**, 7556–7561 (1998).
- Smyth, M. J. *et al.* Perforin-mediated cytotoxicity is critical for surveillance of spontaneous lymphoma. *J. Exp. Med.* **192**, 755–760 (2000).
- Kim, S., Iizuka, K., Aguilu, H. L., Weissman, I. L. & Yokoyama, W. M. *In vivo* natural killer cell activities revealed by natural killer cell-deficient mice. *Proc. Natl Acad. Sci. USA* **97**, 2731–2736 (2000).
- Vance, R. E., Tanamachi, D. M., Hanke, T. & Raulet, D. H. Cloning of a mouse homolog of CD94 extends the family of C-type lectins on murine natural killer cells. *Eur. J. Immunol.* **27**, 3236–3241 (1997).
- Ho, E. L. *et al.* Murine Nkg2d and Cd94 are clustered within the natural killer complex and are expressed independently in natural killer cells. *Proc. Natl Acad. Sci. USA* **95**, 6320–6325 (1998).
- Wu, J. *et al.* An activating immunoreceptor complex formed by NKG2D and DAP10. *Science* **285**, 730–732 (1999).
- Diefenbach, A., Jamieson, A. M., Liu, S. D., Shastri, N. & Raulet, D. H. Ligands for the murine NKG2D receptor: expression by tumor cells and activation of NK cells and macrophages. *Nature Immunol.* **1**, 119–126 (2000).
- Bauer, S. *et al.* Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA. *Science* **285**, 727–729 (1999).
- Cerwenka, A. *et al.* Retinoic acid early inducible genes define a ligand family for the activating NKG2D receptor in mice. *Immunity* **12**, 721–727 (2000).
- Cosman, D. *et al.* ULBP3, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. *Immunity* **14**, 123–133 (2001).
- Nomura, M. *et al.* Genomic structures and characterization of Rael family members encoding GPI-anchored cell surface proteins and expressed predominantly in embryonic mouse brain. *J. Biochem.* **120**, 987–995 (1996).
- Groh, V. *et al.* Broad tumor-associated expression and recognition by tumor-derived gamma delta T cells of MICA and MICB. *Proc. Natl Acad. Sci. USA* **96**, 6879–6884 (1999).
- Gays, F. *et al.* The mouse tumor cell lines EL4 and RMA display mosaic expression of NK-related and certain other surface molecules and appear to have a common origin. *J. Immunol.* **164**, 5094–5102 (2000).
- van Hall, T. *et al.* Identification of a novel tumor-specific CTL epitope presented by RMA, EL-4, and MBL-2 lymphomas reveals their common origin. *J. Immunol.* **165**, 869–877 (2000).
- Huang, A. *et al.* Role of bone marrow-derived cells in presenting MHC class I-restricted tumor antigens. *Science* **264**, 961–965 (1994).
- Gruftman, P. & Karre, K. Innate and adaptive immunity to tumors: IL-12 is required for optimal responses. *Eur. J. Immunol.* **30**, 1088–1093 (2000).
- Dranoff, G. *et al.* Vaccination with irradiated tumor cells engineered to secrete murine granulocyte-macrophage colony-stimulating factor stimulates potent, specific, and long-lasting anti-tumor immunity. *Proc. Natl Acad. Sci. USA* **90**, 3539–3543 (1993).
- Groh, V. *et al.* Costimulation of CD8  $\alpha\beta$  T cells by NKG2D via engagement by MIC induced on virus-infected cells. *Nature Immunol.* **2**, 255–260 (2001).
- Glas, R. *et al.* Recruitment and activation of natural killer (NK) cells *in vivo* determined by the target cell phenotype: an adaptive component of NK cell-mediated responses. *J. Exp. Med.* **191**, 129–138 (2000).
- Pende, D. *et al.* Identification and molecular characterization of Nkp30, a novel triggering receptor involved in natural cytotoxicity mediated by human natural killer cells. *J. Exp. Med.* **190**, 1505–1516 (1999).
- Moretta, L. *et al.* Activating receptors and coreceptors involved in human natural killer cell-mediated cytotoxicity. *Annu. Rev. Immunol.* **19**, 197–223 (2001).
- Medzhitov, R. & Janeway, C. A. Jr. Innate immunity: the virtues of a nonclonal system of recognition. *Cell* **91**, 295–298 (1997).
- Hart, I. R. The selection and characterization of an invasive variant of the B16 melanoma. *Am. J. Pathol.* **97**, 587–600 (1979).
- van Elsland, A., Hurwitz, A. A. & Allison, J. P. Combination immunotherapy of B16 melanoma using anti-cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) and granulocyte/macrophage colony-stimulating factor (GM-CSF)-producing vaccines induces rejection of subcutaneous and metastatic tumors accompanied by autoimmune depigmentation. *J. Exp. Med.* **190**, 355–366 (1999).
- Liao, N., Bix, M., Zijlstra, M., Jaenisch, R. & Raulet, D. MHC class I deficiency: susceptibility to natural killer (NK) cells and impaired NK activity. *Science* **253**, 199–202 (1991).
- Murali-Krishna, K. *et al.* Counting antigen-specific CD8 T cells: a reevaluation of bystander activation during viral infection. *Immunity* **8**, 177–187 (1998).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

## Acknowledgements

We thank C. W. McMahon, and N. Fernandez for critical reading of the manuscript; N. Fernandez for PK136 antibody; J. Egen for help with the retroviral transfection system; H. Nolla for cell sorting; J. Hsia and C. White for technical assistance; the members of the Raulet laboratory for discussions; and J. P. Allison, M. Fasso, Y. H. Chien and B. Sha for their gift of reagents and cell lines. This work was supported by grants from the National Institute of Health (D.H.R.). A.D. is a Physician Postdoctoral Fellow of the Howard Hughes Medical Institute.

Correspondence and request for materials should be addressed to D.H.R. (e-mail: [raulet@uclink4.berkeley.edu](mailto:raulet@uclink4.berkeley.edu)).

# S-Nitrosothiols signal the ventilatory response to hypoxia

Andrew J. Lipton\*, Michael A. Johnson†, Timothy Macdonald†, Michael W. Lieberman‡, David Gozal\*§ & Benjamin Gaston§||

\*Kosair Children's Hospital Research Institute, Departments of Pediatrics, Pharmacology and Toxicology, University of Louisville, Louisville, Kentucky 40202, USA

†Department of Chemistry, University of Virginia, Charlottesville, Virginia 22906, USA

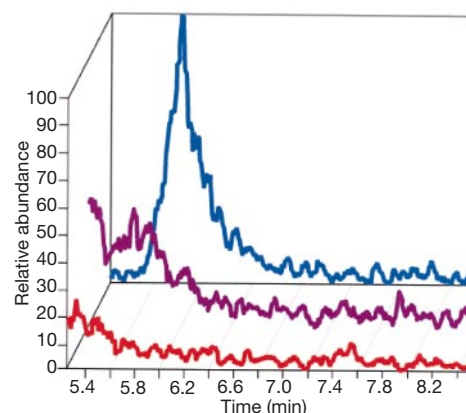
‡Department of Pathology and Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030, USA

||Department of Pediatrics, Division of Respiratory Medicine, University of Virginia School of Medicine, Charlottesville, Virginia 22908, USA

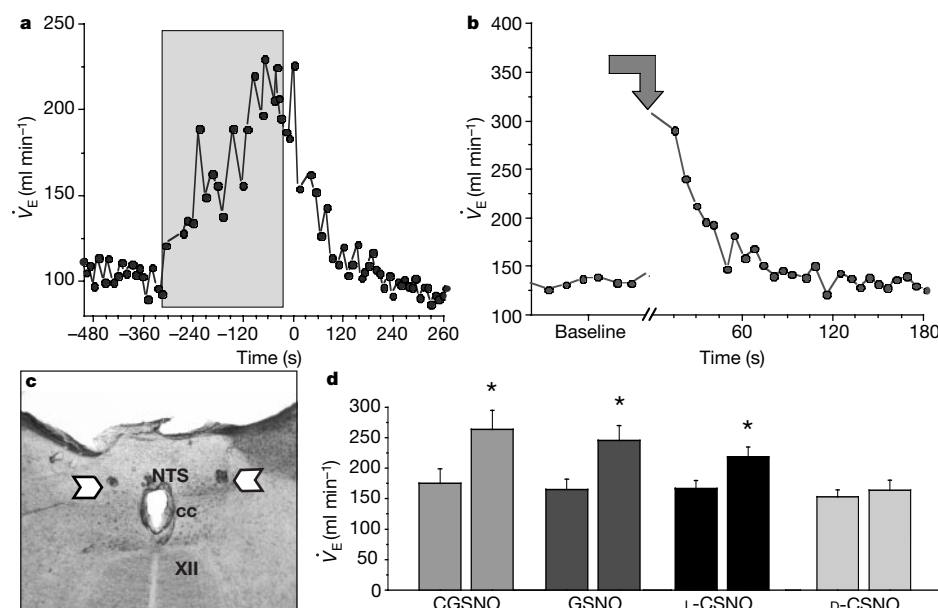
§These authors contributed equally to the work

Increased ventilation in response to hypoxia has been appreciated for over a century<sup>1</sup>, but the biochemistry underlying this response remains poorly understood. Here we define a pathway in which increased minute ventilation ( $\dot{V}_E$ ) is signalled by deoxyhaemoglobin-derived S-nitrosothiols (SNOs). Specifically, we demonstrate that S-nitrosocysteine glycine (CGSNO) and S-nitroso-L-cysteine (L-CSNO)—but not S-nitroso-D-cysteine (D-CSNO)—reproduce the ventilatory effects of hypoxia at the level of the nucleus tractus solitarius (NTS). We show that plasma from deoxygenated, but not from oxygenated, blood produces the ventilatory effect of both SNOs and hypoxia. Further, this activity is mediated by S-nitrosoglutathione (GSNO), and GSNO activation by  $\gamma$ -glutamyl transpeptidase ( $\gamma$ -GT) is required. The normal response to hypoxia is impaired in a knockout mouse lacking  $\gamma$ -GT. These observations suggest that S-nitrosothiol biochemistry is of central importance to the regulation of breathing.

Failure to increase minute ventilation ( $\dot{V}_E$ ) in response to hypoxia may be fatal. In particular, newborn mammals have



**Figure 1** Glutathione reacts with deoxygenated but not oxygenated blood to form GSNO. Liquid chromatography/mass spectrometry was performed on plasma from blood that was reacted with 400  $\mu$ M GSH. GSNO was eluted at 5.82 min. The GSNO peak in the deoxygenated blood-derived plasma (blue) was attenuated after ultraviolet treatment (violet) and is not evident in the oxygenated blood-derived plasma (red). Positive identification of the GSNO peak was determined by co-elution of the sample with a GSNO standard. The retention time and both mass spectrometry and mass spectrometry fragmentation pattern of the endogenous species were identical to exogenous GSNO plus the endogenous species. These observations demonstrate that deoxygenated blood reacts with reduced thiol species to form ultraviolet-labile SNO species, whereas oxygenated blood does not.



**Figure 2** Ventilatory effects of SNOs. **a**,  $\dot{V}_E$  during (shaded) and following a short period of hypoxia. **b**, Injection of 1 nmol CGSNO into the NTS resulted in a marked increase in  $\dot{V}_E$  (injection indicated by an arrow) with onset and decay characteristics identical to those observed during short exposure of the whole animal to hypoxia and return to normoxia. **c**, Neuronal tissue section showing NTS. Arrows indicate injection site. cc, central canal; XII, hypoglossal nucleus. Because conscious animals were used, baseline  $\dot{V}_E$  varied considerably; however, a substantial increase was observed with each L-SNO isomer. **d**, All L-SNO isomers caused increases in  $\dot{V}_E$  (change from baseline for CGSNO: asterisk,  $P < 0.001$ ,  $n = 10$ ; GSNO: asterisk,  $P < 0.0001$ ,  $n = 14$ ; L-CSNO: asterisk,  $P < 0.0001$ ,  $n = 20$ ), whereas D-CSNO was without effect ( $P = \text{NS}$ ;  $n = 20$ ).

marked respiratory instability during hypoxia, as do patients with processes such as obstructive sleep apnoea and anatomic brainstem abnormalities<sup>2</sup>. The mechanisms by which hypoxic stimuli are processed are poorly understood. However, it is known that  $\dot{V}_E$  increases linearly with decreasing oxyhaemoglobin saturation ( $\sim 0.6 \text{ l min}^{-1} (\% \text{ sat})^{-1}$  in health) and that its regulation involves input to brainstem areas such as the NTS, which are rich in nitric oxide synthase (NOS)<sup>3,4</sup>.

GSNO is present in micromolar concentrations in the rat brain<sup>5</sup>. It may be formed when NOS is activated in neuronal and other tissue, perhaps through a secondary reaction<sup>6</sup>, and also from transnitrosation reactions involving the globin  $\beta 93$  S-NO bond allowed during the conformational change of the haemoglobin tetramer during blood deoxygenation<sup>7</sup>. This latter transfer may be facilitated at the erythrocyte membrane by anion-exchange protein 1 (ref. 8). GSNO, in turn, helps to signal adaptations to decreasing oxygen availability<sup>7-9</sup>, including microvascular recruitment—a prominent feature of the circulatory response of the brainstem to hypoxia<sup>10</sup>. Indeed, venous plasma SNO levels are higher relative to arterial values in hypoxic than in normal newborn humans<sup>11</sup>. Consistent with these observations, we found—using both mass spectrometry (Fig. 1) and reductive chemiluminescence (arterial to venous ratio is  $0.26 \pm 0.027$ ;  $n = 4$ )—higher plasma levels of GSNO formed in venous blood that was exposed to erythrocytic and brainstem concentrations (400  $\mu\text{M}$ ) of reduced glutathione (GSH)<sup>8,12</sup> under argon ( $p_{\text{O}_2} \sim 25$  torr) compared with GSH-exposed arterial blood in room air ( $p_{\text{O}_2} \sim 125$  torr). These data demonstrate that deoxygenated whole blood forms SNOs more effectively from corresponding reduced thiols compared with oxygenated blood.

The L- and D-isomers of CSNO release NO at the same rate in rat brain, but only the L-isomer depresses heart rate and blood pressure when injected into the NTS<sup>13,14</sup>. An interaction with a receptor and stereoselective catabolism to form locally active NO have both been proposed to account for neuronal SNO effects, and both may be important<sup>13-15</sup>. Because NOS activation and blood deoxygenation each may result in SNO formation and increased ventilation, we

proposed that SNOs could also stimulate respiratory centres of the NTS to increase  $\dot{V}_E$ .

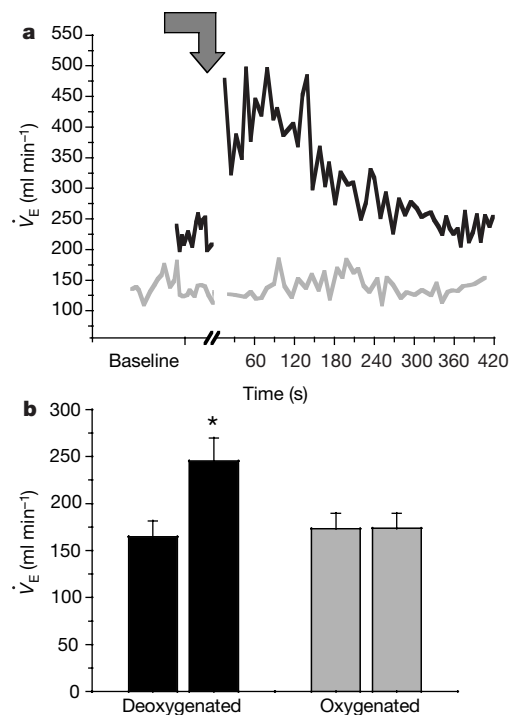
To test this hypothesis we began by examining the ability of endogenous SNOs to increase  $\dot{V}_E$  at the level of the NTS in freely behaving, conscious rats using whole-body plethysmography. CSNO, GSNO and CGSNO (1 nmol each) caused equivalent increases in  $\dot{V}_E$ , whereas D-CSNO had no effect (Fig. 2; dose threshold for L-SNOs is 0.1 pmol). The exogenous NO donor, S-nitroso-N-acetyl-L-penicillamine, had similar but reduced effects (not shown). The L- and D-isomers of CSNO decayed at identical rates in rat brainstem homogenates ( $26\% \text{ min}^{-1} \mu\text{g}^{-1}$  protein each;  $P = \text{not significant (NS)}$ ). Furthermore, neither excess 8-bromo cyclic GMP nor glutathione had any effect on  $\dot{V}_E$  ( $n = 3$ ;  $P = \text{NS}$ ).

Next, we studied whether a low-mass fraction (less than a relative molecular mass of 10,000 ( $M_r$  10K)) derived from deoxygenated blood (Fig. 1) would similarly increase  $\dot{V}_E$ . This fraction reproduced precisely the effect of GSNO, L-CSNO and CGSNO, whereas the fraction from oxygenated blood was without effect (Fig. 3). As expected, ultraviolet photolysis of the deoxygenated, blood-derived fraction—which causes homolytic cleavage of the SNO bond and liberation of free NO (Fig. 1)—completely eliminated its effect on  $\dot{V}_E$  ( $n = 3$ ). These observations suggest that SNOs arising during blood deoxygenation can signal an increase in  $\dot{V}_E$ .

We then studied the normal physiological response to hypoxia in relation to the SNO effect on  $\dot{V}_E$ . Exposure to a 10% oxygen environment resulted in an increase in  $\dot{V}_E$  identical to that produced by L-isomers of SNOs administered to the NTS (Fig. 2). The decay characteristics for the recovery of  $\dot{V}_E$  after injection of GSNO were identical to those for recovery from hypoxia (Fig. 2). This recovery is characterized by a return of  $\dot{V}_E$  to baseline over approximately 3 min after return to normoxia<sup>16</sup>. Taken together, these observations demonstrate that SNOs duplicate the physiological response of exposure to, and recovery from, hypoxia.

Endothelial transport and targeted neuronal biological activity of GSNO may depend, under certain circumstances, on biochemical modifications such as cleavage by  $\gamma$ -GT to form CGSNO<sup>9,17</sup>. We

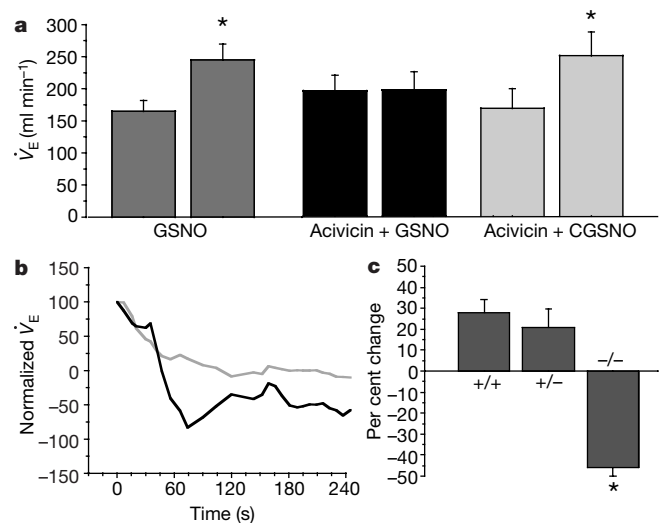




**Figure 3** The low-mass fraction from deoxygenated blood signals an increase in  $\dot{V}_E$  at the level of the NTS. **a**, Microinjection into the NTS of conscious rats of the GSH-derived fraction from deoxygenated blood (black line; see Fig. 1) stimulated a  $\dot{V}_E$  increase that was absent when the low-mass fraction from oxygenated blood was injected (grey line). **b**, The differences in  $\dot{V}_E$  before and after injection of deoxygenated fractions (black;  $n = 14$ ) were highly reproducible (asterisk,  $P < 0.0001$ ), whereas oxygenated fractions (grey;  $n = 12$ ) had no effect ( $P = \text{NS}$ ).

found that NTS pre-treatment with the  $\gamma$ -GT inhibitor, acivicin<sup>19</sup>, attenuated GSNO-mediated increases in  $\dot{V}_E$ , and that CGSNO overcame this inhibition (Fig. 4). Moreover, normal ventilatory offset (that is, the gradual return to baseline of  $\dot{V}_E$  after recovery from hypoxia that prevents apnoea) was inhibited by acivicin. This suggests that GSNO is a critical precursor of CGSNO—and perhaps dipeptidase-derived L-CSNO—in determining hypoxic ventilatory responses *in vivo* (Fig. 4). As expected<sup>17</sup>, we found that deoxygenated blood-derived CGSNO was less stable than GSNO, consistent with a pathway by which GSNO is activated locally by  $\gamma$ -GT in neuronal tissue. Of note, this effect of  $\gamma$ -GT may distinguish respiratory stimuli to the NTS (which are reproduced by GSNO and are  $\gamma$ -GT dependent) from haemodynamic effects of SNOs, which although stereoselective are not reproduced by GSNO<sup>13</sup>. This distinction may have pharmacological implications. Finally, we demonstrated conclusively that  $\gamma$ -GT is required for the normal ventilatory response to hypoxia using a mouse deficient in  $\gamma$ -GT<sup>26</sup>. Homozygous deficient mice had profound attenuation of hypoxic ventilatory recovery (Fig. 4).

These results show that endogenous SNOs act stereoselectively at the level of the NTS to produce the normal ventilatory response seen during hypoxia. This fits well with the proposal that SNOs may serve as signalling molecules between endothelial cells and central and peripheral neurons<sup>12</sup>, as well as recent observations that (1) haemoglobin deoxygenation is associated with an increase in SNO formation and biological activity<sup>7–9</sup>; (2) neuronal tissues contain high levels of SNOs<sup>5</sup>; (3) afferents from peripheral respiratory chemoreceptors project to areas of the NTS rich in NOS, which may produce SNOs<sup>3,4,6,18</sup>; (4) NOS expression by NTS is increased after chronic hypoxia (D.G., unpublished observations); and (5) hypoxic ventilatory responses are attenuated both by NOS inhibition<sup>20,21</sup> and in endothelial NOS knockout mice<sup>22</sup>.



**Figure 4** Effect of  $\gamma$ -GT inhibition or deficiency on the ventilatory effects of SNOs. **a**,  $\dot{V}_E$  increases stimulated by microinjection of 10 nmol GSNO were abolished after pre-treatment with the  $\gamma$ -GT inhibitor acivicin (7.5 nmol;  $P < 0.0001$ ;  $n = 8$ ). CGSNO (10 nmol) stimulated  $\dot{V}_E$  increases that were not modified by acivicin ( $P = \text{NS}$ ;  $n = 6$ ). **b**, Breath-by-breath  $\dot{V}_E$  immediately after cessation of hypoxia following vehicle (grey line) or acivicin (black line) injection. Analysis of ventilatory decay during hypoxia offset and after microinjection of GSNO in 6 rats reveals similar kinetics for both stimulus schemes. Pre-treatment with acivicin was associated with marked acceleration of ventilatory decay after hypoxia. **c**, Hypoxic ventilatory response in  $\gamma$ -GT-deficient mice (+/+, wild type; +/-, heterozygotes; --/,  $\gamma$ -GT-deficient). Lowest  $\dot{V}_E$  during the 30 s after cessation of hypoxia is expressed as per cent change from pre-hypoxia baseline ( $P < 0.0001$ ;  $n = 8$  each group).

We have also shown that a compound derived from deoxygenated, but not oxygenated, blood reproduces the ventilatory effect of hypoxia. This biological activity is physiologically identical to both exogenous SNOs and hypoxia itself, and was identified as GSNO by mass spectrometry. Both this SNO activity and hypoxia require  $\gamma$ -GT for normal signalling. Additional responses to hypoxia attributed to the transfer of nitrogen oxides from deoxygenated blood include vasodilatation to maintain oxygen delivery<sup>8</sup> and increases in endothelial gene expression mediated by hypoxia-inducible factor-1 (ref. 9). We suggest that S-nitrosothiol signalling is of central importance in the normal response to hypoxia, and that this pathway will provide targets for the development of new treatments for apnoea. □

## Methods

### Measurement of $\dot{V}_E$

A dual cannula (22G; Plastics One) was implanted close to the nucleus of the solitary tract according to standard stereotaxic coordinates (−14.0 mm bregma, 0.5 mm off midline, 7.0 mm depth; see ref. 23 figure 74) through a hole drilled into the occipital skull of male Sprague–Dawley rats (~250 g; pentobarbital anaesthesia). We confirmed placement histologically after protocol completion. After recovery for 48 h (return to normal feeding and sleep/waking patterns), breath-by-breath ventilation was measured using the barometric method previously described<sup>21</sup> after simultaneous, bilateral 0.1- $\mu$ l injections in the freely behaving animal (see Fig. 2). Protocols were approved by the institutional Animal Use and Care Committee.

### Plasma preparation

Oxygenated blood drawn in an airtight syringe with GSH and EDTA (final concentrations 400  $\mu$ M) by left ventricular puncture (rat) or peripherally (human) was maintained in 21% oxygen. Venous blood was reacted identically and was transferred to a hypoxia chamber under argon. We measured oxygen tension electrochemically (Chirion). Samples underwent centrifugation (3,000g for 5 min) followed by ultrafiltration (10K; Millipore), separation and selective ultraviolet photolysis (Jelight PS-3000-30; Laguna-Hills).

## Preparation of and assay for S-nitrosothiols

SNOs were prepared by acid nitrosation<sup>24</sup>, titrated to pH 7.4 and maintained in EDTA in the dark at -80 °C until use to prevent decomposition. SNOs were detected by reduction/chemiluminescence as previously described<sup>25</sup>. Briefly, samples were injected into 100 µM CuCl<sub>2</sub>, 1 mM cysteine (pH 6; 50 °C) and purged with helium (grade 5; BOC Gases). We measured evolved NO by chemiluminescence (Sievers).

## Mass spectrometry

Samples eluted isocratically (90% of 0.1% formic acid, 10% methanol, 1.5 ml min<sup>-1</sup>) over a Waters Symmetry C18 column (7.8 × 150 mm) were collected, lyophilized and reconstituted. Purified samples were injected onto a Waters C18 microbore column (1.0 × 150 mm) and analysed by electrospray ionization mass spectrometry using a Finnigan LCQ Duo system. GSNO cations were monitored by selective ion monitoring at a mass to charge ratio (*m/z*) of 336.9. For mass spectrometry and mass spectrometry fragmentation experiments, GSNO cations were dissociated in the ion trap, and the fragments were monitored within a *m/z* range of 90–350.

Received 19 April; accepted 25 June 2001.

- Haldane, J. & Smith, J. L. The physiological effects of air vitiated by respiration. *J. Pathol. Bacteriol.* **1**, 168–186 (1892).
- Bigard, G. E. & Neubauer, J. A. In *Regulation of Breathing* 2nd edn (eds Dempsey, J. A. & Pack, A. I.) (Marcel Dekker, New York, 1995).
- Lin, L. H., Emson, P. C. & Talman, W. T. Apposition of neuronal elements containing nitric oxide synthase and glutamate in the nucleus tractus solitarius of rat: a confocal microscopic analysis. *Neuroscience* **96**, 341–350 (2000).
- Gozal, D. *et al.* Nitric oxide modulates ventilatory responses to hypoxia in conscious developing rats. *Am. J. Respir. Crit. Care Med.* **155**, 1755–1762 (1997).
- Kluge, I., Gutteck-Amsler, U., Zollinger, M. & Do, K. Q. S-Nitrosoglutathione in rat cerebellum: identification and quantification by liquid chromatography-mass spectrometry. *J. Neurochem.* **69**, 2599–2607 (1997).
- Schmidt, H. *et al.* No NO from NO synthase. *Proc. Natl Acad. Sci. USA* **93**, 14492–14497 (1996).
- Jia, L., Bonaventura, C., Bonaventura, J. & Stamler, J. S. S-Nitrosohaemoglobin: a dynamic activity of blood involved in vascular control. *Nature* **380**, 221–226 (1996).
- Pawloski, J. R., Hess, D. T. & Stamler, J. S. Export by red blood cells of nitric oxide bioactivity. *Nature* **409**, 622–626 (2001).
- Palmer, L. A., Gaston, B. & Johns, R. A. Normoxic stabilization of hypoxia inducible factor 1 expression and activity by nitrosation donors. *Mol. Pharmacol.* **58**, 1197–1203 (2000).
- Anwar, M., Kissen, I. & Weiss, H. R. Effect of chemodenervation on the cerebral vascular response and microvascular response to hypoxia. *Circ. Res.* **67**, 1365–1373 (1990).
- Gaston, B. *et al.* Umbilical arterial S-nitrosothiols in stressed newborns: role in perinatal circulatory transition. *Biochem. Biophys. Res. Commun.* **253**, 899–901 (1998).
- Chieh, C. C. & Rauhala, P. The redox pathway of S-nitrosoglutathione, glutathione and nitric oxide in cell to neuron communications. *Free Rad. Res.* **31**, 641–650 (1999).
- Ohta, H., Bates, J., Lewis, S. & Talman, W. Actions of S-nitrosocysteine in the nucleus tractus solitarius are unrelated to release of nitric oxide. *Brain Res.* **746**, 98–104 (1997).
- Davisson, R. L., Travis, M. D., Bates, J. N., Johnson, A. K. & Lewis, S. J. Stereoselective actions of S-nitrosocysteine in central nervous system of conscious rats. *Am. J. Physiol.* **272**, H2361–H2368 (1997).
- Lipton, S. A. *et al.* A redox-based mechanism for the neuroprotective and neurodestructive effects of nitric oxide and related nitroso-compounds. *Nature* **364**, 626–632 (1993).
- Menendez, A. A., Nuckton, T. J., Torres, J. E. & Gozal, D. Short-term potentiation of ventilation after different levels of hypoxia. *J. Appl. Physiol.* **86**, 1478–1482 (1999).
- Askew, S. C., Butler, A. R., Flitney, F. W., Kemp, G. D. & Megson, I. L. Chemical mechanisms underlying the vasodilator and platelet anti-aggregating properties of S-nitroso-N-acetyl-DL-penicillamine and S-nitrosoglutathione. *Biorgan. Med. Chem.* **3**, 1–9 (1995).
- Choi, Y. *et al.* Molecular basis of NMDA receptor-coupled ion channel modulation by S-nitrosylation. *Nature Neurosci.* **3**, 15–21 (2000).
- Smith, T. K., Ikeda, Y., Fujii, J., Taniguchi, N. & Meister, A. Different sites of acivicin binding and inactivation of γ-glutamyl-transpeptidases. *Proc. Natl Acad. Sci. USA* **92**, 2360–2364 (1995).
- Torres, J. E., Kreisman, N. R. & Gozal, D. Nitric oxide modulates *in vitro* intrinsic optical signal and neural activity in the nucleus tractus solitarius of the rat. *Neurosci. Lett.* **232**, 175–178 (1997).
- Gozal, D., Torres, J. E., Gozal, Y. M. & Littwin, S. M. Effect of nitric oxide synthase inhibition on cardiorespiratory responses in the conscious rat. *J. Appl. Physiol.* **81**, 2068–2077 (1996).
- Kline, D., Yang, T., Premkumar, D., Thomas, A. & Prabhakar, N. Blunted respiratory responses to hypoxia in mutant mice deficient in nitric oxide synthase-3. *J. Appl. Physiol.* **88**, 1496–1508 (2000).
- Paxinos, G. & Watson, C. *The Rat Brain in Stereotaxic Coordinates* (Academic, New York, 1986).
- Stamler, J. S. & Feelisch, M. In *Methods in Nitric Oxide Research* (eds Stamler, J. S. & Feelisch, M.) 521–539 (Wiley, Chichester, 1996).
- Fang, K., Ragsdale, N. V., Carey, R. M., Macdonald, T. & Gaston, B. Reductive assays for S-nitrosothiols: implications for measurements in biological systems. *Biochem. Biophys. Res. Commun.* **252**, 535–540 (1998).
- Lieberman, M. W. *et al.* Growth retardation and cysteine deficiency in γ-glutamyl transpeptidase-deficient mice. *Proc. Natl Acad. Sci. USA* **93**, 7923–7926.

## Acknowledgements

This work was supported by the NIH/NHLBI and the Commonwealth of Kentucky Research Challenge Trust Fund.

Correspondence and requests for material should be addressed to B.G. (e-mail: bmg3g@virginia.edu).

# The SH2/SH3 adaptor Grb4 transduces B-ephrin reverse signals

Chad A. Cowan & Mark Henkemeyer

Center for Developmental Biology and Kent Waldrep Foundation Center for Basic Research on Nerve Growth and Regeneration, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9133, USA

**Bidirectional signals mediated by membrane-anchored ephrins and Eph receptor tyrosine kinases have important functions in cell–cell recognition events, including those that occur during axon pathfinding<sup>1,2</sup> and hindbrain segmentation<sup>3,4</sup>. The reverse signal that is transduced into B-ephrin-expressing cells is thought to involve tyrosine phosphorylation of the signal's short, conserved carboxy-terminal cytoplasmic domain<sup>5,6</sup>. The Src-homology-2 (SH2) domain proteins that associate with activated tyrosine-phosphorylated B-subclass ephrins have not been identified, nor has a defined cellular response to reverse signals been described. Here we show that the SH2/SH3 domain adaptor protein Grb4 binds to the cytoplasmic domain of B ephrins in a phosphotyrosine-dependent manner. In response to B-ephrin reverse signalling, cells increase FAK catalytic activity, redistribute paxillin, lose focal adhesions, round up, and disassemble F-actin-containing stress fibres. These cellular responses can be blocked in a dominant-negative fashion by expression of the isolated Grb4 SH2 domain. The Grb4 SH3 domains bind a unique set of other proteins that are implicated in cytoskeletal regulation, including the Cbl-associated protein (CAP/ponsin), the Abl-interacting protein-1 (Abi-1), dynamin, PAK1, hnRNPK and axin. These data provide a biochemical pathway whereby cytoskeletal regulators are recruited to Eph–ephrin bidirectional signalling complexes.**

Eph receptors and ephrins exhibit a unique mode of signalling in that each molecule can function as both a ligand capable of sending signals and a receptor capable of transducing signals. There are three B-subclass ephrins, encoded by different genes, which have a hydrophobic transmembrane segment and a short, conserved cytoplasmic tail containing a number of tyrosine residues. When cells that express B ephrins are co-cultured with cells that express their cognate EphB receptors, both molecules become activated as measured by the induction of phosphorylation of cytoplasmic tyrosine residues<sup>5</sup>. Experiments where B-ephrin-expressing cells were challenged with an aggregated soluble form of the EphB2 extracellular domain (EphB2-Fc) resulted in B-ephrin tyrosine phosphorylation<sup>5,6</sup>, thereby providing further evidence of a function for ephrins as receptors and Eph receptors as ligands. Eph receptors possess an intrinsic tyrosine kinase catalytic domain that auto-phosphorylates when activated, whereas the cytoplasmic tails of B ephrins are thought to be phosphorylated by other kinases, such as Src or Fyn. The idea that ephrins and Eph receptors can each function as both a receptor and a ligand is consistent with biological data indicating non-cell-autonomous functions for Eph family receptors<sup>1,2,7</sup> and cell-autonomous functions for ephrins<sup>3,4,8</sup>. Although a reasonable amount of information concerning the biochemical 'forward' signal transduced by the Eph receptor is known, the identity of proteins that transduce the 'reverse' signal after activation of B ephrins is not known.

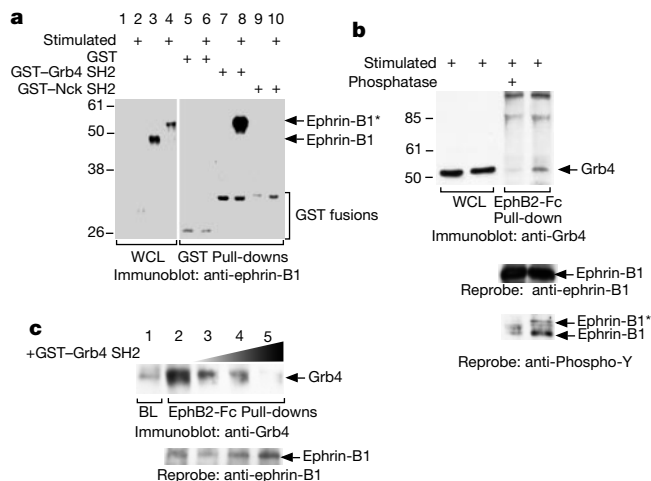
A modified yeast two-hybrid screen, which allows for the detection of protein–protein interactions involving phosphotyrosine, was used to identify partners of the ephrin-B1 cytoplasmic domain (see Methods). A total of 19 positive clones was recovered from a murine neonatal brain complementary DNA library that interacted with ephrin-B1 in a phosphotyrosine-dependent manner. All 19 clones contained overlapping sequences of the same adaptor pro-

tein, Grb4 (also known as Nck-2 and Nck $\beta$ )<sup>9–11</sup>, and all clones encoded its intact SH2 domain. Although it is possible that additional SH2 domains will be able to recognize and bind the ephrin-B1 cytoplasmic tail, no other SH2-containing proteins were recovered in our screen. A region of ephrin-B1 that can bind to Grb4 was narrowed down to a 22-amino-acid peptide (CPHYEKVSG-DYGHVPYIVQEMP), which is identical between ephrin-B1 and ephrin-B2 and contains three of the five conserved tyrosine residues. Other tests in yeast determined that Grb4 could interact with all three B-ephrin cytoplasmic tails and that both the murine and human Grb4 SH2 domains (96% identical) were capable of binding to the ephrins (data not shown).

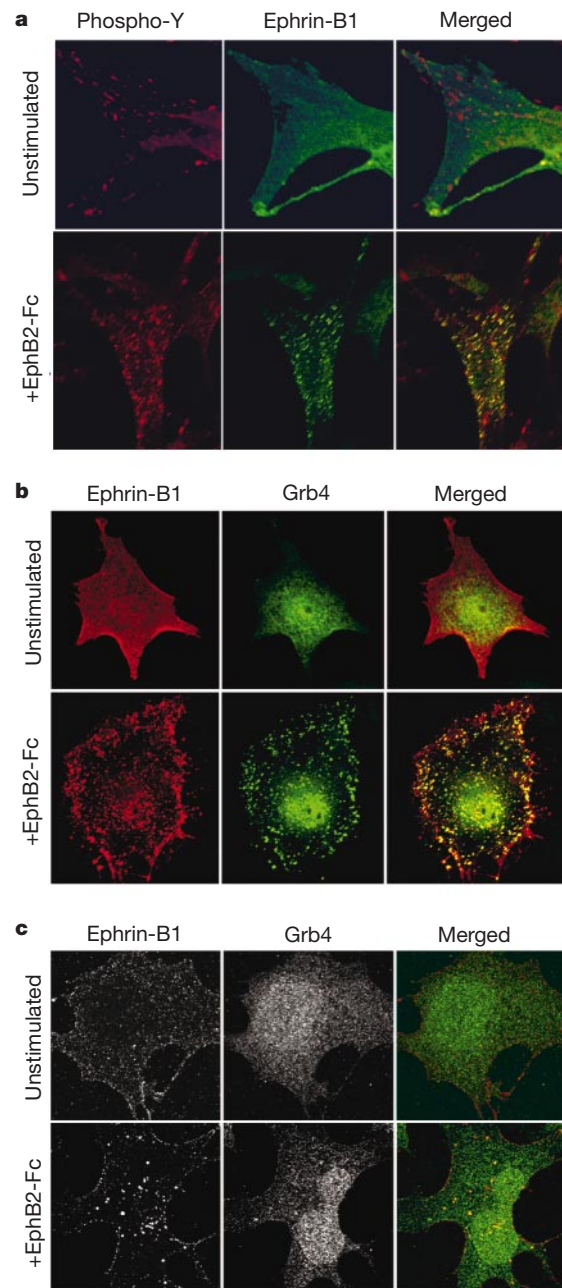
The closest known relatives to Grb4 are another mammalian adaptor protein, Nck, and the *Drosophila* protein Dreadlocks (DOCK). All three proteins share the same SH3–SH3–SH3–SH2 domain organization and, notably, studies implicate the DOCK protein in the pathfinding of certain axons in the fly<sup>12</sup> and the Grb4/Nck proteins in actin dynamics<sup>13,14</sup>. Although DOCK has been reported to be the *Drosophila* homologue of Nck, an alignment shows that it is equally related to both Nck and Grb4 (44% identical and 65% similar). The Grb4 and Nck SH2 domains share 85% identity, and a direct test in yeast determined that the Nck SH2 domain did not interact with the tyrosine-phosphorylated ephrin-B1 cytoplasmic domain (not shown), indicating that they are not functionally interchangeable.

Glutathione *S*-transferase (GST) pull-down experiments were used to verify the interaction of Grb4 with ephrin-B1 in mammalian cells. Neuronal NG108 cells engineered to express full-length ephrin-B1 were stimulated to induce tyrosine phosphorylation, and proteins that bound to the Grb4 SH2 domain were immunoblotted with anti-ephrin-B1 antibodies (Fig. 1a). In these GST pull-down experiments, the Grb4 SH2 domain showed very strong binding only to tyrosine-phosphorylated ephrin-B1, whereas the Nck SH2 domain showed little or no binding. We next investigated

whether ephrin-B1 and Grb4 could be co-precipitated from neuronal CHP100 cells, which express these proteins endogenously. To induce reverse signalling, CHP100 cells were exposed to soluble EphB2-Fc that had been pre-clustered with anti-human immunoglobulin- $\gamma$  (IgG). Whole-cell lysates were then mixed with the same EphB2-Fc reagent to precipitate ephrin-B1 and any associated proteins. In these experiments, Grb4 was found to co-precipitate



**Figure 1** The Grb4 SH2 domain binds tyrosine-phosphorylated ephrin-B1. **a**, Mock- (lanes 1 and 2) and ephrin-B1-transfected (lanes 3–10) NG108 cells were stimulated (+) and whole cell lysates (WCL) were mixed with GST fusion proteins. Bound proteins were immunoblotted, revealing that phosphorylated ephrin-B1 (asterisk) was specifically precipitated by the Grb4 SH2 domain. Relative molecular masses (in thousands) of protein standards are indicated on the left. **b**, CHP100 cells were stimulated for 2 h and then ephrin-B1-associated proteins were immunoblotted, revealing Grb4 bound to tyrosine-phosphorylated ephrin-B1 (asterisk). The association was diminished by alkaline-phosphatase treatment of the precipitates. **c**, Brain lysates (BL) were incubated with EphB2-Fc to precipitate B-ephrins and bound proteins were immunoblotted, revealing that Grb4 co-precipitated with ephrin-B1 (lane 2). The association was competed away by increasing amounts of a bacterially produced Grb4 SH2 domain (lanes 2–5).



**Figure 2** Grb4 associates with activated ephrin-B1 in live cells. **a**, **b**, Serum-starved BHK cells transfected to express ephrin-B1 (**a**) or ephrin-B1 and HA-tagged Grb4 (**b**) were exposed to pre-clustered EphB2-Fc for 45 min, fixed, and then processed for immunofluorescence with the indicated antibodies. In **a**, ephrin-B1 became localized to distinct spots on the plasma membrane (green), which contain phosphotyrosine (red). The merged images show overlap of both signals (yellow). In unstimulated cells the phosphotyrosine signal is localized to focal adhesions. **b**, After activation and clustering of ephrin-B1 (red), Grb4 (green) co-localized to the same spots on the plasma membrane. **c**, After activation of reverse signalling in CHP100 cells, endogenous Grb4 co-localized to spots of endogenous ephrin-B1.



with ephrin-B1 in a phosphotyrosine-dependent manner (Fig. 1b). In similar EphB2-Fc pull-down experiments, Grb4 co-precipitated with ephrin-B1 in protein lysates obtained from adult mouse brains (Fig. 1c).

We also investigated whether the recruitment of Grb4 to ephrin-B1 could be visualized in live cells after activation of reverse

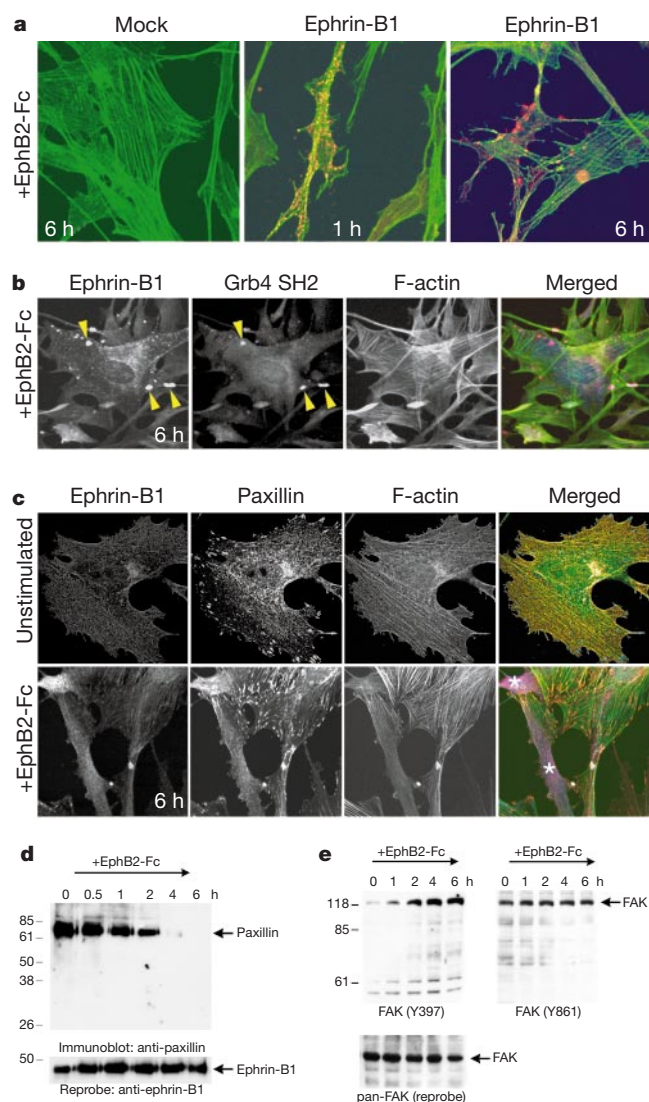
signalling. To accomplish this, BHK fibroblast cells that were engineered to express both ephrin-B1 and a haemagglutinin (HA)-tagged version of Grb4 were stimulated with soluble, pre-clustered EphB2-Fc. This resulted in ephrin-B1 becoming localized to distinct aggregates in the plasma membrane that contained a strong phosphotyrosine signal (Fig. 2a) and that co-localized Grb4 (Fig. 2b). Endogenous Grb4 was also found to co-localize to spots of activated, endogenous ephrin-B1 in CHP100 cells after exposure to pre-clustered EphB2-Fc (Fig. 2c). Additional experiments determined that Nck does not co-localize to ephrin-B1 in response to reverse signalling, and that Grb4 recruitment requires the ephrin-B1 cytoplasmic domain (see Supplementary Information). Together, these experiments provide strong evidence that Grb4 associates with the tyrosine-phosphorylated cytoplasmic tail of B-ephrins after the activation of reverse signalling.

Bidirectional signals that are transduced by Eph receptors and ephrins are generally thought to control cell and axon growth-cone movement by repulsion, presumably by inducing changes in the cytoskeleton. To directly examine the cytoskeleton after reverse signalling, serum-starved, transiently transfected BHK cells were exposed to pre-clustered EphB2-Fc and then double labelled for ephrin-B1 protein and phalloidin to visualize F-actin-containing stress fibres. As shown in Fig. 3a, stimulation of ephrin-B1 reverse signalling resulted in a marked loss of polymerized actin as revealed by the disassembly of stress fibres, which normally span the length of the cell. In response to long periods of exposure to EphB2-Fc, ephrin-B1-expressing cells were observed to have rounded up and shrunk as though adhesive contacts had been lost, while adjacent non-transfected cells were unaffected. Coexpression of ephrin-B1 and the isolated SH2 domain of Grb4 (Fig. 3b), but not that of Nck (see Supplementary Information), blocked the ability of ephrin-B1 to induce these cellular responses in a dominant-negative fashion.

In addition to the loss of stress fibres and changes in cell adhesion, the focal adhesion marker paxillin was found to re-distribute from the cell membrane to the cytoplasm in response to reverse signalling (Fig. 3c, d). Furthermore, an analysis of FAK tyrosine-phosphorylation levels using a series of antibodies directed against its six main phosphotyrosine residues (397, 407, 576, 577, 861 and 925) indicated that only one amino acid (397) is altered in response to reverse signalling—becoming hyperphosphorylated (Fig. 3e and data not shown). This particular tyrosine is within the kinase domain of FAK and its phosphorylation has been previously shown to be important for its catalytic activity<sup>31</sup>. Consistent with the cytoskeletal changes observed after B-ephrin reverse signalling, increases in FAK catalytic activity have been implicated in the disassembly of focal adhesions.

To gain an insight into how Grb4 may contribute to reverse signalling and induce both the disassembly of focal adhesions and the collapse of stress fibres, we identified proteins from the neonatal mouse brain cDNA library that interact in yeast with a bait composed of the three Grb4 SH3 domains. SH3 domains form high-affinity protein–protein interactions with molecules that have proline-rich segments<sup>15</sup>, and indeed most of the positive clones from our screen were found to encode known proteins with at least one stretch of poly-proline residues, including the expected PXXP motif. We did recover *bona fide* targets of SH3 domains, as two of the identified proteins have been shown to bind SH3 domains—dynamitin<sup>16</sup> and heterogeneous nuclear ribonucleoprotein K (hnRNPK)<sup>17,18</sup>. We also identified multiple clones encoding three different, known signalling proteins not previously implicated to exhibit interactions with SH3 domains: the Abl interacting protein-1 (Abi-1), the Cbl-associated protein (CAP/ponsin), and a regulator of Wnt/ $\beta$ -catenin signalling, axin.

GST pull-down experiments verified the interactions between Grb4 and the proteins that were identified in the yeast two-hybrid screen (Fig. 4a). Mimicking the results in yeast, GST–Abi-1, GST–axin, GST–CAP, GST–dynamitin and GST–hnRNPK fusion proteins precipitated full-length Grb4 from NG108 cell lysates. Although it

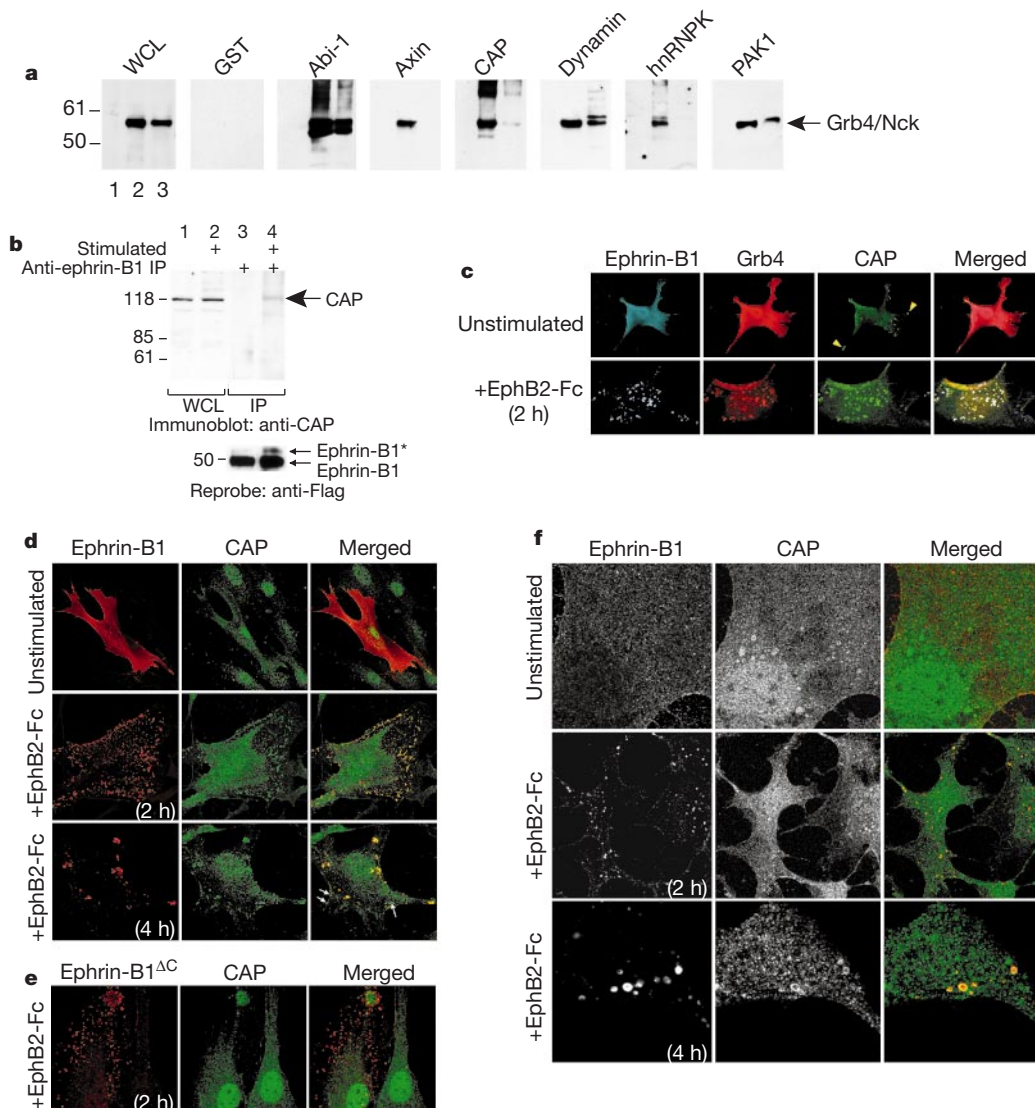


**Figure 3** Activation of ephrin-B1 reverse signalling leads to loss of stress fibres and disassembly of focal adhesions. **a**, BHK cells transiently transfected with ephrin-B1 were stimulated, fixed and then labelled to detect polymerized actin (green) and ephrin-B1 (red). Very little F-actin and no stress fibres can be detected in the ephrin-B1-expressing cell shown at the 6-h time point, whereas adjacent, non-expressing and mock-transfected cells appear normal. **b**, The isolated Grb4 SH2 domain co-localized to activated ephrin-B1 (yellow arrowheads) and blocked the disassembly of F-actin stress fibres. **c**, After exposure to pre-clustered EphB2-Fc, untransfected BHK cells remain flat and show no change in paxillin localization. This protein is lost from focal adhesions in cells expressing ephrin-B1 (marked with asterisks). Note that ephrin-B1-expressing cells have lost attachments with the substratum, have rounded up, and have shrunk. **d**, **e**, Stable ephrin-B1-expressing NG108 cells were stimulated for the indicated time points. Relative molecular masses (in thousands) of protein standards are indicated on the left. In **d**, cells were lysed and membrane-associated proteins were immunoblotted, revealing a loss of paxillin from the membrane after reverse signalling. The reprobe demonstrates that other membrane-associated proteins were not lost at the later time points. In **e**, stimulated cells were lysed and total proteins were immunoblotted with FAK phosphotyrosine-specific antibodies, revealing an increase in the phosphorylation of Tyr 397.

was not identified in our yeast screen, a GST-PAK1 fusion protein was also able to associate with Grb4 in these pull-down assays. Abi-1, dynamin and PAK1 interacted with both Grb4 and Nck, whereas axin, CAP and hnRNP $\kappa$  showed specificity for Grb4.

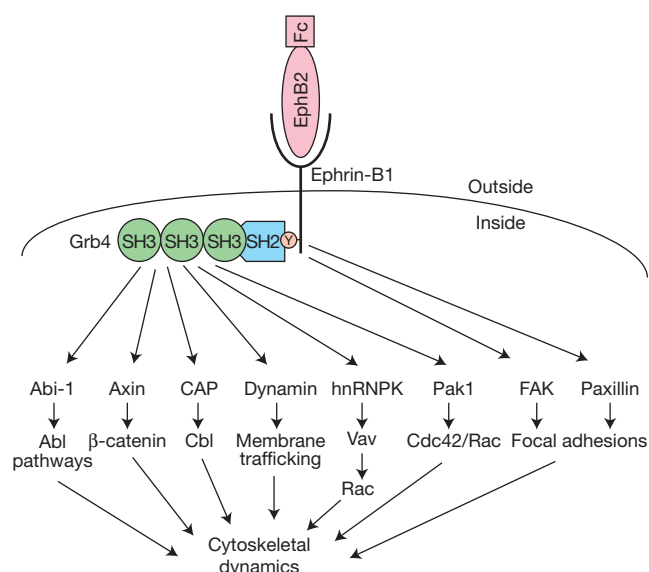
To further characterize these protein-protein interactions, we focused on CAP, as this molecule has been implicated to function at focal adhesions<sup>19,20</sup>. Biochemical studies demonstrated that ephrin-B1 and endogenous CAP can be co-immunoprecipitated from NG108 cells after activation of reverse signalling (Fig. 4b). CAP contains only one PXXP motif (PQQP beginning at residue 604) and an *in vitro* filter-binding assay indicates that the SH3.1 and/or SH3.3 domains of Grb4 are able to recognize this sequence (see

Supplementary Information). Stimulation of ephrin-B1, Grb4 and CAP triple-expressing BHK cells with pre-clustered EphB2-Fc resulted in a redistribution of Grb4 and CAP from focal adhesions to activated spots of ephrin-B1 (Fig. 4c). In additional experiments with BHK cells, endogenous CAP was found to be recruited to spots of activated ephrin-B1 (Fig. 4d), and that co-localization requires the ephrin-B1 cytoplasmic tail (Fig. 4e). Furthermore, coexpression of the isolated Grb4 SH2 domain, but not that of Nck, severely diminished the ability of endogenous CAP to be recruited to ephrin-B1 in a dominant-negative fashion (see Supplementary Information). Finally, using CHP100 cells that endogenously express ephrin-B1, Grb4 and CAP, the recruitment of CAP to spots of ephrin-B1 were



**Figure 4** Grb4 SH3 domains bind cytoskeletal regulators. **a**, Protein lysates from mock (lane 1), HA-tagged Grb4 (lane 2) and HA-tagged Nck (lane 3) transiently transfected NG108 cells were incubated with indicated GST fusion proteins. Bound proteins were immunoblotted with an anti-HA antibody. **b**, After stimulation for 2 h, Flag-tagged ephrin-B1 was immunoprecipitated from NG108 cells and the complexes immunoblotted for endogenous CAP. This treatment revealed an association only in activated cells that contain tyrosine-phosphorylated ephrin-B1 (asterisk). Relative molecular masses (in thousands) of protein standards are indicated on the left. **c**, BHK cells transfected with ephrin-B1, HA-tagged Grb4 and Flag-tagged CAP were exposed to pre-clustered EphB2-Fc and then triple labelled to detect all three proteins. This revealed a co-localization of Grb4 and CAP to spots of activated ephrin-B1. In unstimulated cells,

transfected CAP is localized to focal adhesions (yellow arrowheads). **d**, BHK cells transiently transfected with ephrin-B1 were stimulated and then double labelled to detect ephrin-B1 and endogenous CAP. Note that the ephrin-B1/CAP signalling centres become larger during long stimulations of reverse signalling, and that CAP that is not co-localized to ephrin-B1 is detected in a punctate pattern throughout the cell, along stress fibres and focal adhesions (arrows), and in the nucleus. **e**, Although still able to aggregate to spots in the membrane after exposure to pre-clustered EphB2-Fc, a truncated version of ephrin-B1 lacking most of its cytoplasmic domain did not recruit endogenous CAP. **f**, CHP100 neuronal cells were stimulated, leading to the recruitment of endogenous CAP to spots of activated, endogenous ephrin-B1.



**Figure 5** Summary of proposed B-ephrin signal transduction pathway. Activation of reverse signalling after interactions with EphB2 extracellular domains leads to clustering and tyrosine phosphorylation of the cytoplasmic tail of ephrin-B1, and a concomitant recruitment of Grb4 and its SH3-binding partners. Additionally, ephrin-B1 reverse signalling alters the subcellular localization and/or tyrosine-phosphorylation levels of other proteins involved in cytoskeletal regulation, including paxillin and FAK. Although reverse signalling involves multimerization of the components, only one of each molecule is shown in the diagram for simplicity.

readily observed after the activation of reverse signalling (Fig. 4f). Taken together, these results indicate that, in response to reverse signalling, clustered tyrosine-phosphorylated ephrin-B1 forms a high-affinity binding site for the Grb4 SH2 domain, and that the SH3 domains of this protein piggy-backs CAP and other SH3-binding proteins to sites of activated B-ephrins.

The bidirectional signals transduced by ephrins and Eph receptors are important in a diverse array of cell–cell recognition events, and the reverse signal transduced by B ephrins is thought to lead to axonal repulsion<sup>1,2</sup> and inhibition of cell mixing<sup>3,4</sup>. Both of these cellular responses may result from localized disassembly of the cytoskeleton. We show here that ephrin-B1 reverse signalling can lead to cell rounding, loss of stress fibres, a redistribution of paxillin away from focal adhesions, and an increased phosphorylation of Tyr 397 in FAK. All of these changes are consistent with the purported repellent roles for B-ephrin reverse signalling. Clustered tyrosine-phosphorylated ephrin-B1 forms a high-affinity binding site for the Grb4 SH2 domain, and blocking this docking site can prevent the transduction of reverse signals into the cell. The Grb4 SH3 domains bind a number of other proteins, many of which have been implicated to have roles in cytoskeletal regulation. CAP has been proposed to regulate actin stress fibres and focal adhesions, and is also localized to adherens junctions<sup>19,20</sup>. Notably, CAP has three SH3 domains that mediate associations with other cytoskeletal proteins, including Cbl, the F-actin-binding protein vinculin, AF6 (afadin), and FAK<sup>19,21</sup>. Many of the other identified Grb4 SH3-binding proteins are also implicated in cytoskeletal regulation, including PAK1 kinase, hnRNPK, axin and Abi-1. hnRNPK has been shown to interact with Vav, an exchange factor for Rac<sup>17,18</sup>, whereas PAK1 has roles in regulating Cdc42 and Rac. Axin is known to regulate the stability of β-catenin and control Wnt signalling<sup>22–24</sup>, and is also implicated in axonal remodelling<sup>25</sup> and in cell adhesion by regulating adherens junctions<sup>26</sup>. Furthermore, Armadillo, the β-catenin-related protein in *Drosophila*, has been shown to have a function in regulating the assembly of the axon scaffold in the

central nervous system, and this function may involve the Abl tyrosine kinase<sup>27</sup>. In an additional connection to Abl, we identified Abi-1 as a potential partner for the Grb4 SH3 domains. Abi-1 has its own SH3 domain that binds Abl directly<sup>28,29</sup>.

We have determined that the Grb4 SH2 domain can associate with the cytoplasmic domain of tyrosine-phosphorylated B ephrins, and we have identified a number of interacting proteins capable of binding the Grb4 SH3 domains that may have a role, along with other proteins, in propagating the reverse signal (Fig. 5). Our data provide evidence that Grb4 is involved in transducing reverse signals into B-ephrin-expressing cells and, given the biological importance of reverse signalling, this interaction may have functional significance by regulating the cytoskeleton during cell adhesion, cell migration and axonal pathfinding. □

## Methods

### Yeast two-hybrid screens

A bait composed of LexA fused to the murine ephrin-B1 cytoplasmic domain (residues 246–345) was coexpressed in yeast with the PDGF receptor tyrosine kinase domain<sup>30</sup>. From a total of  $3.96 \times 10^7$  transformants screened with a neonatal mouse brain cDNA library, 96 HIS<sup>+</sup> colonies were recovered that encoded proteins capable of interacting with the phosphorylated ephrin. Nineteen contained in-frame fusions of Grb4 starting at either residue 202 (17 clones) or 245 (2 clones). The other 77 positive colonies contained inserts encoding PDZ domain proteins, which were not found to interact with ephrin-B1 in a phosphotyrosine-dependent manner (C.A.C. and M.H., unpublished data). Directed screening was done with murine ephrin-B1 (residues 313–346 and 313–334), ephrin-B2 (residues 304–337), ephrin-B3 (residues 308–341), Grb4 (residues 287–380) and human GRB4 (residues 286–381) and NCK (residues 282–387). In all tests, transformants were plated on synthetic medium lacking tryptophan, leucine and histidine, supplemented with 25, 50, 75 and 100 mM 3-aminotriazole, and assayed for growth and β-galactosidase activity.

The human GRB4 SH3 domains (residues 1–263) were used to screen the mouse neonatal brain library in the MaV203 strain of yeast. Out of  $4 \times 10^7$  total transformants, 98 HIS<sup>+</sup> and URA3<sup>+</sup> positive clones were recovered and sequenced, 48 of which are presented here and fall into five classes. One positive clone encoded Abi-1 starting at residue 103; 15 encoded axin, all starting at residue 653; five encoded CAP, starting at residue 165 (2), 215 (1) or 323 (2); 21 encoded dynamin, starting at residue 486 (1), 544 (1), 553 (7), 585 (1), 600 (9) or 630 (2); and six encoded hnRNPK, starting at either residue 46 (2) or 197 (4).

### Mammalian expression

A cDNA encoding full-length murine ephrin-B1 was cloned into pCDNA3.1 (Invitrogen). Flag- or Myc-tagged epitopes were introduced either eight amino acids extracellular (between 228–229) or intracellular (between 270–271) of the hydrophobic ephrin-B1 transmembrane domain. The various untagged and tagged versions of ephrin-B1 were used interchangeably without any noticeable differences. The truncated ephrin-B1<sup>ΔC</sup> introduces a stop codon after residue 312. The isolated Grb4 SH2 domain (residues 202–380) was cloned into pCDNA3.1 with a Myc tag at the amino terminus. The isolated human NCK SH2 domain (197–377) was cloned into pCDNA3.1His/Xpress (Invitrogen). HA-tagged human GRB4 and NCK expression vectors were from W. Li, and Flag-tagged CAP expression vector was from C. Baumann and A. Saltiel.

### Transfections and cell stimulation

NG108 cells were transiently transfected with 5–10 μg DNA using calcium phosphate. To produce ephrin-B1-expressing stable cell lines, NG108 cells were transfected with 5 μg DNA and clonal lines were selected in 700 μg ml<sup>−1</sup> G418 (Gibco/BRL), expanded, and maintained in 400 μg ml<sup>−1</sup> G418. BHK cells (strain CCL10) were transiently transfected with 0.75–5 μg DNA using Lipofectamine Plus (Gibco). Cells were stimulated with 5–10 μg ml<sup>−1</sup> pre-clustered EphB2-Fc (R&D Systems) at 37 °C as described<sup>5</sup>.

### Biochemistry

GST fusion proteins were generated by cloning the following murine coding regions into pGEX: Grb4 (residues 287–380), Abi-1 (residues 103–395), axin (residues 653–992), CAP (residues 165–684), dynamin (residues 553–861) and hnRNPK (residues 197–465). pGEX-PAK1 (residues 1–231) was from J. Frost and M. Cobb and pGEX-Nck (residues 282–377) was from T. Pawson. Protein lysates, GST pull-down experiments, EphB2-Fc pull-downs and immunoprecipitations were carried out as described<sup>1,5</sup>. Proteins were resolved on SDS–PAGE gels, transferred to nitrocellulose, and immunoblotted with the indicated antibodies.

### Membrane fractionation

A stable ephrin-B1-expressing NG108 cell line was plated on 15-cm Nunclon dishes, serum-starved for 12 h, and then subjected to a time course of stimulation with 10 μg ml<sup>−1</sup> of pre-clustered EphB2-Fc. Plasma membranes prepared in detergent-free lysis buffer were pelleted by ultracentrifugation at 100,000g and the membrane-associated proteins were immunoblotted with the indicated antibodies.



## Immunofluorescence

CHP100 cells or transfected BHK cells were plated on glass coverslips (Baxter) and stimulated with 5–10  $\mu\text{g ml}^{-1}$  pre-clustered EphB2-Fc for the desired time points. Cells were fixed and then processed according to standard procedures. Images were acquired with a confocal microscope (Leica DM DRB or Zeiss LSM 510) and processed with Adobe Photoshop. See supplementary Information for detailed methods.

Received 22 May; accepted 17 July 2001.

- Henkemeyer, M. *et al.* Nuk controls pathfinding of commissural axons in the mammalian central nervous system. *Cell* **86**, 35–46 (1996).
- Birgbauer, E., Cowan, C. A., Sretavan, D. W. & Henkemeyer, M. Kinase independent function of EphB receptors in retinal axon pathfinding to the optic disc from dorsal but not ventral retina. *Development* **127**, 1231–1241 (2000).
- Mellitzer, G., Xu, Q. & Wilkinson, D. G. Eph receptors and ephrins restrict cell intermingling and communication. *Nature* **400**, 77–81 (1999).
- Xu, Q., Mellitzer, G., Robinson, V. & Wilkinson, D. G. *In vivo* cell sorting in complementary segmental domains mediated by Eph receptors and ephrins. *Nature* **399**, 267–271 (1999).
- Holland, S. J. *et al.* Bidirectional signaling through the Eph family receptor Nuk and its transmembrane ligands. *Nature* **383**, 722–725 (1996).
- Bruckner, K., Pasquale, E. B. & Klein, R. Tyrosine phosphorylation of transmembrane ligands for Eph receptors. *Science* **275**, 1640–1643 (1997).
- George, S. E. *et al.* The VAB-1 Eph receptor tyrosine kinase functions in neural and epithelial morphogenesis in *C. elegans*. *Cell* **92**, 633–643 (1998).
- Davy, A. *et al.* Compartmentalized signaling by GPI-anchored ephrin-A5 requires the Fyn tyrosine kinase to regulate cellular adhesion. *Genes Dev.* **13**, 3125–3135 (1999).
- Margolis, B. *et al.* High-efficiency expression/cloning of epidermal growth factor-receptor-binding proteins with Src homology 2 domains. *Proc. Natl Acad. Sci. USA* **89**, 8894–8898 (1992).
- Chen, M. *et al.* Identification of Nck family genes, chromosomal localization, expression, and signaling specificity. *J. Biol. Chem.* **273**, 25171–25178 (1998).
- Tu, Y., Li, F. & Wu, C. Nck-2, a novel Src homology2/3-containing adaptor protein that interacts with the LIM-only protein PINCH and components of growth factor receptor kinase-signaling pathways. *Mol. Biol. Cell* **9**, 3367–3382 (1998).
- Garrity, P. A. *et al.* *Drosophila* photoreceptor axon guidance and targeting requires the deadlocks SH2/SH3 adaptor protein. *Cell* **85**, 639–650 (1996).
- Lu, W., Katz, S., Gupta, R. & Mayer, B. J. Activation of Pak by membrane localization mediated by an SH3 domain from the adaptor protein Nck. *Curr. Biol.* **7**, 85–94 (1997).
- Chen, M., She, H., Kim, A., Woodley, D. T. & Li, W. Nck-beta adapter regulates actin polymerization in NIH 3T3 fibroblasts in response to platelet-derived growth factor bb. *Mol. Cell. Biol.* **20**, 7867–7880 (2000).
- Ren, R., Mayer, B. J., Cicchetti, P. & Baltimore, D. Identification of a ten-amino acid proline-rich SH3 binding site. *Science* **259**, 1157–1161 (1993).
- Wunderlich, L., Farago, A. & Buday, L. Characterization of interactions of Nck with Sos and dynamin. *Cell Signal* **11**, 25–29 (1999).
- Hobert, O., Jallat, B., Schlessinger, J. & Ullrich, A. Novel signaling pathway suggested by SH3 domain-mediated p95vav/heterogeneous ribonucleoprotein K interaction. *J. Biol. Chem.* **269**, 20225–20228 (1994).
- Bustelo, X. R., Suen, K. L., Michael, W. M., Dreyfuss, G. & Barbacid, M. Association of the vav proto-oncogene product with poly(rC)-specific RNA-binding proteins. *Mol. Cell. Biol.* **15**, 1324–1332 (1995).
- Ribon, V., Herrera, R., Kay, B. K. & Saltiel, A. R. A role for CAP, a novel, multifunctional Src Homology 3 domain-containing protein in formation of actin stress fibers and focal adhesions. *J. Biol. Chem.* **273**, 4073–4080 (1998).
- Mandai, K. *et al.* Ponsin/SH3P12: an 1-afadin- and vinculin-binding protein localized at cell–cell and cell–matrix adherens junctions. *J. Cell Biol.* **144**, 1001–1017 (1999).
- Kurakin, A., Hoffman, N. G. & Kay, B. K. Molecular recognition properties of the C-terminal SH3 domain of the Cbl associated protein, CAP. *J. Pept. Res.* **52**, 331–337 (1998).
- Zeng, L. *et al.* The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulated embryonic axis formation. *Cell* **90**, 181–192 (1997).
- Ikeda, S. *et al.* Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3 $\beta$  and  $\beta$ -catenin and promotes GSK-3 $\beta$ -dependent phosphorylation of  $\beta$ -catenin. *EMBO J.* **17**, 1371–1384 (1998).
- Hamada, F. *et al.* Negative regulation of Wingless signaling by D-axin, a *Drosophila* homolog of axin. *Science* **283**, 1739–1742 (1999).
- Fraklin, L. G., Noordermeer, J. N. & Nusse, R. The *Drosophila* Wnt protein DWnt-3 is a secreted glycoprotein localized on the axon tracts of the embryonic CNS. *Dev. Biol.* **168**, 202–213 (1995).
- Asakura, T. *et al.* Similar and differential behavior between the nectin-afadin-ponsin and cadherin-catenin systems during the formation and disruption of the polarized junctional alignment in epithelial cells. *Genes Cells* **4**, 573–581 (1999).
- Loureiro, J. & Peifer, M. Roles of Armadillo, a *Drosophila* catenin, during central nervous system development. *Curr. Biol.* **8**, 622–632 (1998).
- Dai, Z. & Pendergast, A. M. Abi-2, a novel SH3-containing protein interacts with the c-Abl tyrosine kinase and modulates c-Abl transforming activity. *Genes Dev.* **9**, 2569–2582 (1995).
- Shi, Y., Alin, K. & Goff, S. P. Abl-interactor-1, a novel SH3 protein binding to the carboxy-terminal portion of the Abl protein, suppresses v-abl transforming activity. *Genes Dev.* **9**, 2583–2597 (1995).
- Keegan, K. & Cooper, J. A. Use of the two hybrid system to detect the association of the protein-tyrosine phosphatase, SHPTP2, with another SH2-containing protein, Grb7. *Oncogene* **12**, 1537–1544 (1996).
- Schlaepfer, D. D., Hauck, C. R. & Sieg, D. J. Signaling through focal adhesion kinase. *Prog. Biophys. Mol. Biol.* **71**, 435–478 (1999).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

## Acknowledgements

We thank T. Saxton and T. Pawson for numerous discussions; B. Howell, S. Hollenberg and J. Cooper for two-hybrid vectors and brain library; F. Gertler for CAP peptide spots; K. Womak, X. Cao and S. Kennedy for help with two-hybrid screens; and J. Shay for technical assistance. These studies were made possible by a grant from the Welch Foundation (to M.H.).

Correspondence and requests for materials should be addressed to M.H. (e-mail: [henk@utsw.swmed.edu](mailto:henk@utsw.swmed.edu)).

## CREB regulates hepatic gluconeogenesis through the coactivator PGC-1

Stephan Herzig\*, Fanxin Long\*†, Ulupi S. Jhala\*, Susan Hedrick\*, Rebecca Quinn‡, Anton Bauer§, Dorothea Rudolph§, Gunther Schutz§, Cliff Yoon||, Pere Puigserver||, Bruce Spiegelman|| & Marc Montminy\*

\* Peptide Biology Laboratories, Salk Institute for Biological Studies, 10010 N Torrey Pines Road, La Jolla, California 92037-1002, USA

‡ Joslin Diabetes Center, Boston, Massachusetts 02215, USA

§ Molecular Biology of the Cell I, Deutsches Krebsforschungszentrum Im Neuenheimerfeld 280, D-69120 Heidelberg, Germany

|| Dana-Farber Cancer Center, Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

When mammals fast, glucose homeostasis is achieved by triggering expression of gluconeogenic genes in response to glucagon and glucocorticoids. The pathways act synergistically to induce gluconeogenesis (glucose synthesis), although the underlying mechanism has not been determined<sup>1–4</sup>. Here we show that mice carrying a targeted disruption of the cyclic AMP (cAMP) response element binding (CREB) protein gene, or overexpressing a dominant-negative CREB inhibitor, exhibit fasting hyperglycaemia and reduced expression of gluconeogenic enzymes. CREB was found to induce expression of the gluconeogenic programme through the nuclear receptor coactivator PGC-1, which is shown here to be a direct target for CREB regulation *in vivo*. Overexpression of PGC-1 in CREB-deficient mice restored glucose homeostasis and rescued expression of gluconeogenic genes. In transient assays, PGC-1 potentiated glucocorticoid induction of the gene for phosphoenolpyruvate carboxykinase (PEPCK), the rate-limiting enzyme in gluconeogenesis. PGC-1 promotes cooperativity between cyclic AMP and glucocorticoid signalling pathways during hepatic gluconeogenesis. Fasting hyperglycaemia is strongly correlated with type II diabetes, so our results suggest that the activation of PGC-1 by CREB in liver contributes importantly to the pathogenesis of this disease.

Cyclic AMP regulates the expression of numerous genes through phosphorylation of CREB at Ser 133, mediated by protein kinase A (PKA)<sup>5</sup>. The ability of CREB to stimulate the gluconeogenic PEPCK gene<sup>6,7</sup> prompted us to examine the general importance of this factor for activation of the gluconeogenic programme under fasting conditions. We generated transgenic mice expressing a dominant negative CREB protein, referred to as A-CREB, under control of the liver-specific albumin promoter/enhancer<sup>8</sup>. A-CREB contains the leucine zipper of CREB plus an acidic extension that enhances the

† Present address: Department of Molecular and Cellular Biology, The Biolabs, Harvard University, Cambridge, Massachusetts 02138, USA.

affinity for, and disrupts the DNA-binding activity of, CREB family members (CREB, CREM, ATF-1) but not other bZIP proteins<sup>9,10</sup>. The CRE-binding activity of CREB was reduced by 80% in liver nuclear extracts from A-CREB transgenic mice compared with wild-type littermates, as measured by a gel mobility shift assay (Fig. 1a, compare lanes 2 and 4), despite comparable levels of full-length CREB protein, as shown by a western blot assay (Fig. 1a, bottom). A-CREB mice were profoundly hypoglycaemic (Fig. 1b), and RNA levels for the gluconeogenic enzymes PEPCK and glucose-6-phosphatase (G6Pase) in the liver were lower in these animals than in wild-type littermates (Fig. 1c).

To evaluate acute effects of CREB activity on glucose homeostasis, we prepared an adenovirus expressing A-CREB under the control of the cytomegalo virus (CMV) promoter. Following systemic injection of either A-CREB or control adenovirus, 80–90% of hepatocytes in livers of 6-week-old male mice were infected as determined by visible fluorescence from a co-expressed green fluorescent protein (GFP) marker (not shown). In the fed state, blood glucose levels in A-CREB-infected mice were comparable to those of mice infected with control virus (Fig. 2a;  $1.04 \pm 0.10 \text{ mg ml}^{-1}$  and  $1.35 \pm 0.20 \text{ mg ml}^{-1}$  (mean  $\pm$  s.e.m.) at day 8, respectively). During the fasting period, however, glucose levels were far lower in A-CREB mice than in control injected mice ( $0.65 \pm 0.04 \text{ mg ml}^{-1}$  and  $1.38 \pm 0.05 \text{ mg ml}^{-1}$  at day 8, respectively), demonstrating that CREB activity is required to mobilize glucose via the gluconeogenic pathway (Fig. 2a). Indeed, expression of the PEPCK, G6Pase and pyruvate carboxylase genes was reduced fourfold in livers of A-CREB mice compared with control mice (Fig. 2b). Blood insulin

levels in control and A-CREB mice were comparable in both fed and fasted states (see Supplementary Information Fig. 1).

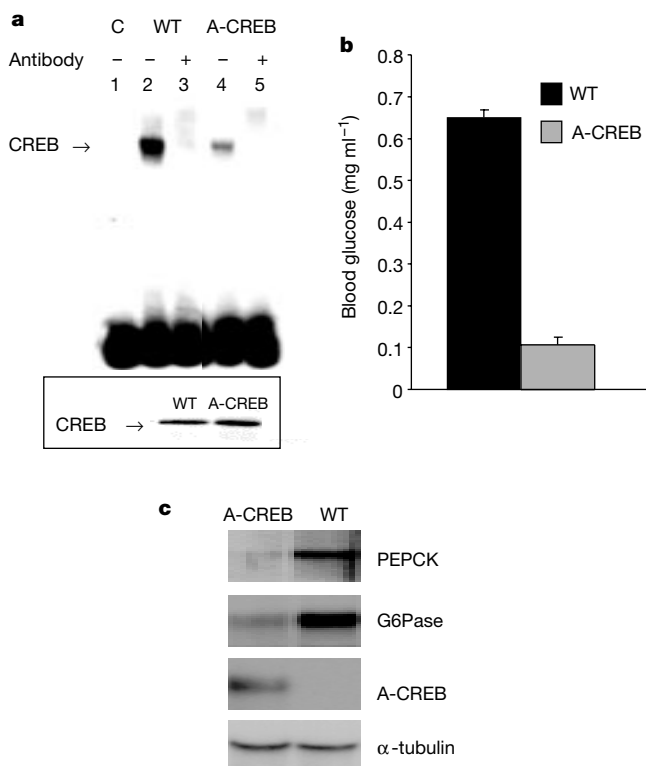
Fasting hyperglycaemia is strongly correlated with insulin resistance in type II diabetic patients<sup>11</sup>. To test the role of CREB in hepatic glucose production in this setting, we employed *db/db* diabetic mice, which display adult-onset diabetes owing to a mutation in the leptin receptor gene<sup>12,13</sup>. Compared with wild-type mice, blood glucose levels in *db/db* mice were significantly elevated in both the fed and fasted states (Fig. 2c;  $2.70 \pm 0.17 \text{ mg ml}^{-1}$  (fed) and  $2.50 \pm 0.16 \text{ mg ml}^{-1}$  (fasted) at day 7). At the same time, PEPCK messenger RNA levels were 2.5-fold higher in fasted *db/db* mice than in wild-type mice (Fig. 2d, compare bars 1 and 3). After injection with A-CREB virus, the blood glucose levels of fasted *db/db* mice returned to normal (Fig. 2c,  $1.22 \pm 0.09 \text{ mg ml}^{-1}$  at day 7) as did PEPCK mRNA levels (Fig. 2d, bar 2). In parallel, plasma insulin levels in fasting *db/db* mice decreased 40% after injection of A-CREB adenovirus (see Supplementary Information Fig. 1).

The reduced expression of G6Pase and pyruvate carboxylase genes in A-CREB-expressing mice is consistent with the hypoglycaemic effects of A-CREB during fasting; disruption of either gene leads to severe hypoglycaemia<sup>14,15</sup>. Remarkably, the G6Pase promoter lacks a discernible CRE<sup>16</sup> and the pyruvate carboxylase gene contains only a weak CRE half-site<sup>17</sup>, suggesting an indirect role for CREB in the regulation of both genes. Expression of PGC-1 has been found to be induced in diabetes and to stimulate gluconeogenic enzymes in cultured hepatocytes<sup>18</sup>, prompting us to examine the role of PGC-1 in mediating CREB signalling under fasting conditions *in vivo*.

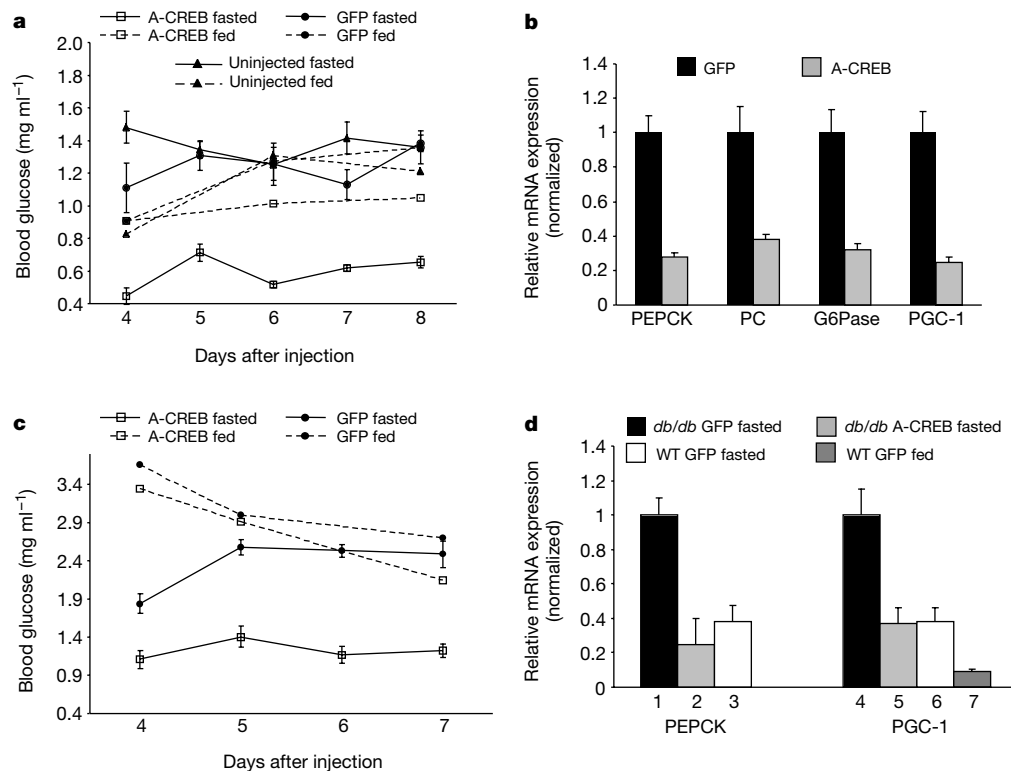
Levels of PGC-1 mRNA in normal adult liver were elevated fourfold after 8 h of food restriction (Fig. 2d, compare bars 6 and 7); and expression of A-CREB blocked PGC-1 induction (Fig. 2b). Consistent with the fasting hyperglycaemia in *db/db* mice, PGC-1 mRNA levels were induced threefold in *db/db* compared with wild-type mice (Fig. 2d, compare bars 4 and 6); overexpression of A-CREB reduced PGC-1 mRNA levels to normal (Fig. 2d, bar 5). Sequence analysis of the mouse PGC-1 promoter revealed a full palindromic consensus CRE centred at –130 (Fig. 3a). To determine whether PGC-1 is indeed a direct target for CREB action *in vivo*, we performed chromatin immunoprecipitation studies using HepG2 hepatoma cells. The PGC-1 promoter was efficiently recovered from immunoprecipitates of CREB but not of control immunoglobulin- $\gamma$  (IgG) (Fig. 3a, compare lanes 3–5). Confirming the specificity of these antisera, no polymerase chain reaction (PCR) product was obtained from CREB immunoprecipitates using control glyceraldehyde-3-phosphate dehydrogenase (GAPDH) primers (Fig. 3a, lanes 3–5).

In transient assays of HepG2 hepatoma cells, PGC-1 promoter activity was induced 3–4-fold by a cAMP agonist (Fig. 3b). Co-transfection of an A-CREB expression vector blocked induction of the PGC-1 promoter by cAMP, demonstrating that PGC-1 is indeed a direct target for CREB activity in hepatocytes (Fig. 3b). Consistent with downregulation of PGC-1 under feeding conditions, PGC-1 reporter activity was blocked in cells treated with insulin (Fig. 3b). The half-maximal dose for inhibition of endogenous PGC-1 mRNA ( $0.1\text{--}0.5 \text{ nM}$ ) was within the physiologic range of insulin concentrations (see Supplementary Information Fig. 2). In keeping with the rapid effects of insulin on hepatic gene expression, PGC-1 mRNA levels were reduced by 70% within 30 min of treatment (see Supplementary Information Fig. 2).

The ability of cAMP to induce a nuclear receptor coactivator in liver prompted us to speculate that PGC-1 promotes cooperativity between cAMP and glucocorticoid pathways during fasting. Towards this end, we co-infected mice with A-CREB and PGC-1 adenovirus constructs. Whereas A-CREB adenovirus alone induced hypoglycaemia in fasted mice, co-infection with PGC-1 adenovirus restored fasting glucose levels to near normal (Fig. 3c) and rescued gluconeogenic gene expression (data not shown). Arguing against

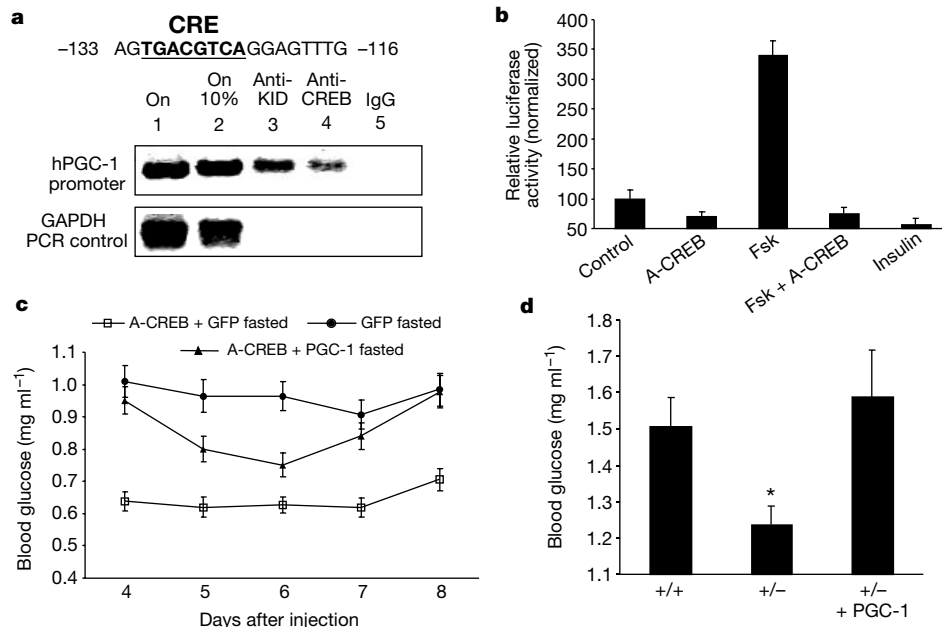


**Figure 1** Transgenic mice expressing a dominant negative CREB inhibitor in liver display profound hypoglycaemia at birth. **a**, Top, gel mobility shift assay of liver nuclear extracts from newborn A-CREB transgenic mice and wild-type (WT) littermates using <sup>32</sup>P-labelled CRE oligonucleotide. Bottom, Western blot of liver nuclear extracts from wild-type and A-CREB transgenic mice using anti-CREB antiserum 244. **b**, Blood glucose measurements of A-CREB transgenic and wild-type littermates at birth before feeding (mean  $\pm$  s.e.m.,  $n = 5$  per group). **c**, Northern blot analysis of PEPCK, G6Pase, A-CREB and tubulin RNAs in livers from wild-type and A-CREB transgenic mice.



**Figure 2** CREB activity is required for glucose homeostasis during fasting. **a**, Blood glucose levels in 6-week-old male mice ( $n = 6$  per group) injected with control (GFP) adenovirus or A-CREB adenovirus, or left uninjected, under fed or fasted (9 h) conditions. **b**, Levels of mRNA for PEPCK, G6Pase and pyruvate carboxylase (PC) in livers ( $n = 3$  per

group) from control (GFP) and A-CREB adenovirus-infected fasting mice. **c**, Blood glucose levels in *db/db* diabetic mice infected with control (GFP) and A-CREB viruses as in **a**. **d**, PEPCK and PGC-1 mRNA levels in fed and fasted mice injected with control (GFP) or A-CREB viruses. Data are means  $\pm$  s.e.m.



**Figure 3** CREB promotes hepatic gluconeogenesis through PGC-1 during fasting. **a**, Chromatin immunoprecipitation assay of HepG2 cells using CREB-specific antisera (anti-KID, anti-CREB). PCR amplification products from reactions with human PGC-1 promoter (-170 to +68) or GAPDH coding region (+586 to +1037) primers. Lanes 1 and 2, output controls. Lanes 3 and 4, PCR products obtained from immunoprecipitates of CREB using antisera specific for anti-KID (amino acids 100–160) and anti-CREB (amino acids 1–341). Lane 5, control immunoprecipitation with non-specific IgG. The PGC-1

promoter CRE element is shown above. **b**, Transient assay of HepG2 cells transfected with PGC-1 luciferase reporter (-170 to +68) and treated with forskolin (Fsk, 10  $\mu$ M), A-CREB expression vector or insulin. **c**, Glucose levels of fasting mice injected with GFP, A-CREB plus GFP, or A-CREB plus PGC-1 adenoviruses. **d**, Glucose levels in fasting CREB<sup>+/-</sup> and wild-type littermates infected with GFP or PGC-1 virus for 8 days. Data are means  $\pm$  s.e.m. Asterisk,  $P < 0.05$  ( $n = 5$ ).



a non-specific effect of PGC-1 on A-CREB expression, levels of the A-CREB inhibitor protein were identical in mice injected with A-CREB plus PGC-1 viruses and in mice injected with A-CREB alone (see Supplementary Information Fig. 3).

To test further the importance of PGC-1 in promoting hepatic gluconeogenesis through CREB, we employed CREB<sup>+/-</sup> mice<sup>19</sup>. CREB<sup>+/-</sup> heterozygotes had lower fasting blood glucose levels than wild-type littermates (Fig. 3d); and injection of PGC-1 adenovirus corrected the relative glucose imbalance in fasted CREB<sup>+/-</sup> mice, revealing the ability of this coactivator to act downstream of CREB to promote hepatic gluconeogenesis (Fig. 3d).

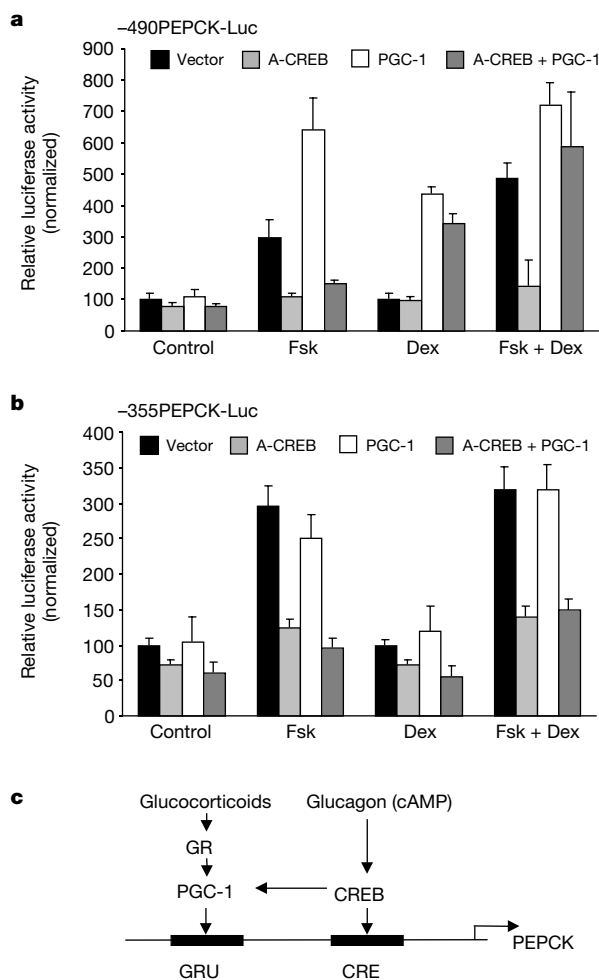
To determine the mechanism by which PGC-1 regulates expression of gluconeogenic genes, we performed transient assays of HepG2 cells using a PEPCK reporter plasmid. Treatment with a cAMP agonist stimulated PEPCK promoter activity 3–4-fold (Fig. 4a); and, consistent with the presence of a consensus CRE sequence in the PEPCK promoter<sup>20</sup>, co-transfection of A-CREB inhibitor plasmid blocked reporter induction (Fig. 4a). Overexpression of PGC-1 potentiated cAMP-dependent induction of a –490 PEPCK promoter construct in HepG2 cells, but had no effect on a

–355 PEPCK reporter lacking the glucocorticoid response unit (GRU; –455 to –340) (Fig. 4b). The GRU consists of binding sites for the glucocorticoid receptor plus additional transcription factors such as HNF-4 and HNF-3 that collectively mediate induction of the PEPCK promoter in response to glucocorticoids and insulin resistance<sup>21,22</sup>.

The ability of PGC-1 to interact with, and potentiate the activity of, the glucocorticoid receptor<sup>23</sup> prompted us to examine the effect of PGC-1 on induction of the PEPCK promoter by glucocorticoids. Dexamethasone treatment or PGC-1 overexpression alone had little effect on PEPCK promoter activity in HepG2 cells (Fig. 4a), but co-transfection of PGC-1 expression vector strongly induced –490 PEPCK promoter activity in response to dexamethasone (Fig. 4a). Supporting a GRU-dependent mechanism of activation, PGC-1 had no effect on a PEPCK reporter construct lacking the GRU (–355 PEPCK) in cells treated with either dexamethasone or cAMP agonist (Fig. 4b).

Our results argue for a two-step model in the activation of gluconeogenic genes by CREB during fasting (Fig. 4c). Ser 133 phosphorylation of CREB in response to catecholamine and glucagon stimulation during the early fasting period induces certain gluconeogenic enzymes (PEPCK) through a direct CREB-mediated effect. Under prolonged fasting, CREB further potentiates gluconeogenic genes (PEPCK, pyruvate carboxylase and G6Pase) by inducing expression of the coactivator PGC-1 in liver. PGC-1 then mediates activation of the gluconeogenic programme in response to glucocorticoid signals. This model is consistent with the observation that the GRU in the PEPCK gene is abnormally activated in type II diabetes, and that treatment with glucocorticoid antagonist reduces blood glucose levels in *db/db* diabetic mice<sup>22</sup>.

The effect of A-CREB on liver gene expression suggests that CREB may constitute an ideal target for therapeutic intervention. Although use of a dominant negative inhibitor such as A-CREB may not be feasible in this regard, small molecules that block CREB phosphorylation or disrupt recruitment of the CREB coactivator CBP (CREB binding protein) may prove effective. Such compounds may be particularly beneficial as adjunctive therapy in lowering fasting blood glucose levels in type II diabetics.



**Figure 4** PGC-1 mediates induction of the PEPCK gene through the GRU. Transient assay of HepG2 hepatoma cells transfected with PEPCK luciferase reporter vectors either (a) containing (–490) or (b) lacking (–355) the GRU. Treatment with forskolin (Fsk, 10  $\mu$ M) and/or dexamethasone (Dex, 10<sup>–7</sup> M) is indicated. The effect of PGC-1 and A-CREB vectors on PEPCK promoter activity is shown. Data are means  $\pm$  s.e.m. c, Model for cooperation between PGC-1 and CREB on the PEPCK promoter. CREB stimulates the PEPCK promoter activity directly through a consensus CRE and indirectly through induction of PGC-1, which induces the GRU through interaction with the glucocorticoid receptor (GR).

## Methods

### Recombinant adenoviruses

Adenoviruses expressing either A-CREB or PGC-1 were generated through homologous recombination between a linearized transfer vector pAD-Track and the adenoviral backbone vector pAD-Easy as described<sup>24</sup>. pAD-A-CREB contained a *NotI/SpeI* fragment from ZEO-A-CREB encoding a Flag-tagged A-CREB polypeptide<sup>25</sup>. pAD-PGC-1 carried the entire murine PGC-1 complementary DNA encoding all but the carboxy-terminal four amino acids. Both viruses co-expressed the GFP transcribed from a second independent CMV promoter to monitor viral infection efficiency. Adenovirus coding for GFP only (pAD-GFP) was used as a control in all experiments. Viruses were purified by the CsCl method and dialysed against PBS buffer containing 10% glycerol before animal injections as described<sup>26</sup>.

### Animal experiments

Male 6-week-old C57Bl6 or *db/db* mice were obtained from The Jackson Laboratory and maintained on a 12-h light–dark cycle with regular unrestricted diet. For virus injections, animals were anaesthetized with Iso-Flurane and a total of 1  $\times$  10<sup>9</sup> plaque-forming units per recombinant virus was administered by a systemic tail vein injection. In each experiment at least six animals received identical treatments. Animals were fasted for 9–18 h overnight with free access to water. Blood glucose was monitored at the end of each fasting period for at least seven consecutive days. Blood glucose was also measured in the non-fasted state immediately before food withdrawal. Blood samples were collected from the tail vein. Liver tissue for RNA and protein isolation was immediately frozen in liquid nitrogen. Additional liver samples were fixed in 10% formaldehyde, sectioned with a cryomicrotome, and investigated for viral infection efficiency by fluorescence microscopy. Studies with CREB knockout mice carrying a targeted disruption of the CREB DNA-binding domain and luciferase zipper region<sup>19</sup> were performed as described above.

### Blood metabolites

Blood glucose values were determined from whole blood using an automatic glucose monitor (One Touch, Lifescan). Plasma insulin levels were determined using a commercial insulin enzyme-linked immunosorbent assay kit (Crystal Chem).

# TaqMan and northern blot analyses

PolyA<sup>+</sup> RNA and total RNA were extracted from homogenized mice livers using the Fast Track 2.0 (Invitrogen) or the RNeasy (Qiagen) kit. RNA samples were treated with DNase I (Promega) and RNA quality was assessed by gel electrophoresis. cDNA was prepared by reverse transcription of 250 ng mRNA or 500 ng total RNA using the Superscript II enzyme and Oligo dT primer (GIBCO BRL). The resulting cDNAs were amplified using the SYBR green PCR kit and an ABI PRISM 7700 Sequence detector (Perkin Elmer). All mRNA expression data from the TaqMan PCR with reverse transcription was normalized to GAPDH expression in the corresponding sample. Northern blot assays were performed as previously described<sup>25</sup>.

# Protein analysis

Protein was extracted from frozen liver samples in SDS-urea-lysis buffer and 20 µg of protein were loaded onto a 12% SDS-polyacrylamide gel and blotted onto nitrocellulose membrane. Western blot assays were performed as previously described<sup>25</sup>.

# Chromatin immunoprecipitation assay

Human hepatoma HepG2 cells were grown to 90% confluence, and the chromatin immunoprecipitation assay was performed as described elsewhere<sup>27</sup>. Specific antibodies against either full-length CREB (253) or the kinase inducible domain (KID) (244) of CREB were used for the immunoprecipitation. Normal rabbit IgG served as a negative control. Precipitated DNA fragments were analysed by PCR amplification using primers directed against the human PGC-1 promoter region, or the GAPDH coding region as negative control.

# Plasmids

Expression plasmids pZEO-A-CREB<sup>25</sup> and pcDNA3-PGC-1<sup>28</sup> have been described previously. The liver-specific A-CREB transgene was constructed by inserting A-CREB cDNA into a liver-specific albumin promoter/enhancer plasmid<sup>8</sup>. To construct luciferase expression plasmids -490PEPCK-Luc and -355PEPCK-Luc BamHI/BglII fragments of the PEPCK promoter region containing 490 or 355 base pairs (bp) of the mouse PEPCK 5'-flanking region, respectively, were cloned into the pGL3 basic luciferase reporter vector (Promega). PGC-1 reporter mPGC-Luc was constructed by inserting a 230-bp fragment of the mouse PGC-1 promoter containing 170 bp of the 5'-flanking region and 68 bp of exon 1 into the pGL3 vector.

# Cell culture and transfection assays

Human hepatoma HepG2 cells were transfected using the Lipofectamine 2000 reagent (GIBCO BRL) according to the manufacturer's instructions (500 ng of indicator plasmid per well). Where indicated, expression plasmids encoding A-CREB (pZeo-A-CREB)<sup>25</sup> or PGC-1 (pcDNA3-PGC-1)<sup>28</sup> were cotransfected (50 ng and 800 ng plasmid per well, respectively). Cotransfections were performed with a constant amount of DNA by adding the empty pcDNA3 vector (Invitrogen). Cells were treated with forskolin (10 µM) and/or dexamethasone (100 nM) or insulin (100 nM) for 14 h. Cell extracts were prepared 48 h after transfection and the luciferase assay was performed as described previously<sup>29</sup>, normalizing to activity from cotransfected Rous sarcoma virus-β-galactosidase expression plasmid (100 ng plasmid per well).

Received 29 May; accepted 25 July 2001.

1. Hanson, R. W. & Reshef, L. Regulation of phosphoenolpyruvate carboxykinase (GTP) gene expression. *Annu. Rev. Biochem.* **66**, 581–611 (1997).
2. Lemaigre, F. P. & Rousseau, G. G. Transcriptional control of genes that regulate glycolysis and gluconeogenesis in adult liver. *Biochem. J.* **303**, 1–14 (1994).
3. Angrand, P. O., Coffinier, C. & Weiss, M. C. Response of the phosphoenolpyruvate carboxykinase gene to glucocorticoids depends on the integrity of the cAMP pathway. *Cell Growth Differ.* **5**, 957–966 (1994).
4. Imai, E., Miner, J. N., Mitchell, J. A., Yamamoto, K. R. & Granner, D. K. Glucocorticoid receptor–cAMP response element-binding protein interaction and the response of the phosphoenolpyruvate carboxykinase gene to glucocorticoids. *J. Biol. Chem.* **268**, 5353–5356 (1993).
5. Gonzalez, G. A. & Montminy, M. R. Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at Serine 133. *Cell* **59**, 675–680 (1989).
6. Liu, J. S., Park, E. A., Gurney, A. L., Roesler, W. J. & Hanson, R. W. Cyclic AMP induction of phosphoenolpyruvate carboxykinase (GTP) gene transcription is mediated by multiple promoter elements. *J. Biol. Chem.* **266**, 19095–19102 (1991).
7. Quinn, P. G. & Granner, D. K. Cyclic AMP-dependent protein kinase regulates transcription of the phosphoenolpyruvate carboxykinase gene but not binding of nuclear factors to the cyclic AMP regulatory element. *Mol. Cell. Biol.* **10**, 3357–3364 (1990).
8. Liu, J. K., Bergman, Y. & Zaret, K. S. The mouse albumin promoter and a distal upstream site are simultaneously DNase I hypersensitive in liver chromatin and bind similar liver-abundant factors in vitro. *Genes Dev.* **2**, 528–541 (1988).
9. Ahn, S. *et al.* A dominant negative inhibitor of CREB reveals that it is a general mediator stimulus-dependent transcription of *c-fos*. *Mol. Cell. Biol.* **18**, 967–977 (1998).
10. Long, F., Schipani, E., Asahara, H., Kronenberg, H. & Montminy, M. The CREB family of activators is required for endochondral bone development. *Development* **128**, 541–550 (2001).
11. Consoli, A. Role of liver in pathophysiology of NIDDM. *Diabetes Care* **15**, 430–441 (1992).
12. Lee, G. H. *et al.* Abnormal splicing of the leptin receptor in diabetic mice. *Nature* **379**, 632–635 (1996).

13. Chen, H. *et al.* Evidence that the diabetes gene encodes the leptin receptor: identification of a mutation in the leptin receptor gene in *db/db* mice. *Cell* **84**, 491–495 (1996).
14. Wexler, I. D. *et al.* Molecular characterization of pyruvate carboxylase deficiency in two consanguineous families. *Pediatr. Res.* **43**, 579–584 (1998).
15. Lei, K. J. *et al.* Glucose-6-phosphatase dependent substrate transport in the glycogen storage disease type-1a mouse. *Nature Genet.* **13**, 203–209 (1996).
16. Schmoll, D. *et al.* Identification of a cAMP response element within the glucose-6-phosphatase hydrolytic subunit gene promoter which is involved in the transcriptional regulation by cAMP and glucocorticoids in H4IIE hepatoma cells. *Biochem. J.* **338**, 457–463 (1999).
17. Jitrapakdee, S., Booker, G. W., Cassidy, A. I. & Wallace, J. C. The rat pyruvate carboxylase gene structure. Alternate promoters generate multiple transcripts with the 5'-end heterogeneity. *J. Biol. Chem.* **272**, 20522–20530 (1997).
18. Yoon, J. C. *et al.* Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature* **413**, 131–138 (2001).
19. Rudolph, D. *et al.* Impaired fetal T cell development and perinatal lethality in mice lacking the cAMP response element binding protein. *Proc. Natl Acad. Sci. USA* **95**, 4481–4486 (1998).
20. Short, J. M., Wynshaw-Boris, A., Short, H. P. & Hanson, R. W. Characterization of the phosphoenolpyruvate carboxykinase (GTP) promoter-regulatory region. II. Identification of cAMP and glucocorticoid regulatory domains. *J. Biol. Chem.* **261**, 9721–9726 (1986).
21. Imai, E. *et al.* Characterization of a complex glucocorticoid response unit in the phosphoenolpyruvate carboxykinase gene. *Mol. Cell. Biol.* **10**, 4712–4719 (1990).
22. Friedman, J. E. *et al.* Phosphoenolpyruvate carboxykinase (GTP) gene transcription and hyperglycemia are regulated by glucocorticoids in genetically obese *db/db* transgenic mice. *J. Biol. Chem.* **272**, 31475–31481 (1997).
23. Knutti, D., Kaul, A. & Kralli, A. A tissue-specific coactivator of steroid receptors, identified in a functional genetic screen. *Mol. Cell. Biol.* **20**, 2411–2422 (2000).
24. He, T. *et al.* A simplified system for generating recombinant adenoviruses. *Proc. Natl Acad. Sci. USA* **95**, 2509–2514 (1998).
25. Michael, L. F., Asahara, H., Shulman, A., Kraus, W. & Montminy, M. The phosphorylation status of a cyclic AMP-responsive activator is modulated via a chromatin-dependent mechanism. *Mol. Cell. Biol.* **20**, 1596–1603 (2000).
26. Becker, T. *et al.* Use of recombinant adenovirus for metabolic engineering of mammalian cells. *Methods Cell Biol.* **43**, 161–189 (1994).
27. Shang, Y., Hu, X., DiRenzo, J., Lazar, M. & Brown, M. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* **103**, 843–852 (2000).
28. Puigserver, P. *et al.* A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* **92**, 829–839 (1998).
29. Nakajima, T. *et al.* RNA helicase A mediates association of CBP with RNA polymerase II. *Cell* **90**, 1107–1112 (1997).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

# Acknowledgements

We thank J. Frangioni for help with adenovirus constructs, C. Arias for help with histology, and K. Suter for performing injections. This work was supported by grants from the National Institutes for Health to M.M. and from the DFG to S.H.

Correspondence and requests for materials should be addressed to M.M. (e-mail: [montminy@salck.edu](mailto:montminy@salck.edu)).

# erratum

# Roles of tumour localization, second signals and cross priming in cytotoxic T-cell induction

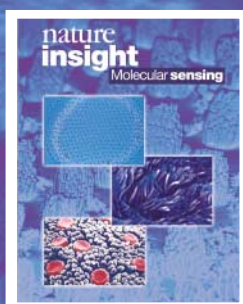
Adrian F. Ochsenbein, Sophie Sierro, Bernhard Odermatt, Marcus Pericin, Urs Karrer, Ian Hermans, Silvio Hemmi, Hans Hengartner & Rolf M. Zinkernagel

Nature **411**, 1058–1064 (2001).

In this Letter, the first name of Ian Hermans was misspelled as 'Jan'. □

# nature insight

## Molecular sensing



### Cover illustrations

Section through *Drosophila* eye (M. Abbey/SPL), and coloured SEMs of ciliated nasal epithelium (BSIP VEM/SPL) and papillae on the tongue (Omikron/SPL). Background: coloured SEM of hair bundles from chicken cochlea (P. G. Gillespie).

### Nature London

The Macmillan Building,  
4 Crinan St,  
London N1 9XW, UK  
Tel +44 207 833 4000  
Fax +44 207 843 4596/7  
e-mail: [nature@nature.com](mailto:nature@nature.com)  
<http://www.nature.com>

### Nature Washington

968 National Press Building,  
529 14th St NW,  
Washington DC 20045, USA  
Tel +1 202 737 2355  
Fax +1 202 628 1609  
e-mail: [nature@nature.com](mailto:nature@nature.com)  
<http://www.nature.com>

### Nature Tokyo

MG Ichigaya Building 5F,  
19-1 Haraikatomachi,  
Shinjuku-ku,  
Tokyo 162-0841, Japan  
Tel +81 3 3267 8751  
Fax +81 3 3267 8746  
e-mail: [nature@naturejpn.com](mailto:nature@naturejpn.com)  
<http://www.naturejpn.com>



**T**he loss of a sense is not life threatening, yet it can severely affect one's quality of life. The first and crucial step in sensory processing — the transduction of stimuli, such as odour, light and sound, into a cellular response — takes place in specialized cells that form an interface between our environments and our nervous systems. Each sense has evolved a transduction mechanism so finely tuned that it is able to discriminate between different stimuli with both speed and sensitivity.

The past few years have seen an explosion in the identification of molecules involved in the different transduction mechanisms. Indeed, this year heralds the tenth anniversary of the discovery of the first odour receptors. These receptors belong to a large family of G-protein-coupled receptors, which amplify signals via intracellular signalling cascades — a mechanism shared by several other senses including vision and taste.

The diversity of signals that our senses must encode is vast. It is remarkable therefore that evolution has repeatedly called upon two ion-channel families to impart such functional diversity. TRP channels were discovered in the fruitfly, where they are involved in the transduction of both light and touch. Another family member, VR1, has a direct role in mammalian detection of noxious heat. Similarly, DEG/ENAC family members are involved in senses ranging from touch in nematodes to mineral taste in mammals. Small wonder, then, that such molecular switches are being engineered for use in commercial biosensor devices.

We are pleased to acknowledge the financial support of NIH Institutes in producing this Insight. As always, *Nature* carries sole responsibility for all editorial content and peer review.

**Lesley Anson Senior Editor and Insight Programme Editor**

**Editor, Nature:** Philip Campbell  
**Insight Editor:** Lesley Anson  
**Editorial Assistant:** Simon Gibbon  
**Production Editor:** Simon Gribbin  
**Art Director:** Majo Xeridat

**Diagrams:** Ann Thomson  
Vicky Askew  
Suzanne Coleman  
**Production Manager:** Yvonne Strong  
**Sponsorship:** Emma Jones

### review articles:

**186 Visual transduction in *Drosophila***  
R. C. Hardie & P. Raghu

**194 Molecular basis of mechanosensory transduction**  
P. G. Gillespie  
& R. G. Walker

**203 Molecular mechanisms of nociception**  
D. Julius  
& A. I. Basbaum

**211 How the olfactory system makes sense of scents**  
S. Firestein

**219 Receptors and transduction in taste**  
B. Lindemann

### progress:

**226 Stochastic sensors inspired by biology**  
H. Bayley & P. S. Cremer

Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Nature Publishing Group (The Macmillan Building, 4 Crinan Street, London N1 9XW). Registered as a newspaper at the British Post Office. Annual subscription for the Americas US\$595 (institutional/corporate), US\$159 (individual making personal payment). Canada residents please add 7% GST (No. 140911595). North and South American orders to: Nature, Subscription Dept, P. O. Box 5055, Brentwood, TN 37024-5055, USA. Other orders to Nature, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Periodicals postage paid at New York, NY 10010-1707, and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by Nature to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of \$12.00 an article (or \$2.00 a page) is paid direct to CCC, 222 Rosewood Drive, Danvers, MA 01923, USA. Identification code for Nature: 0028-0836/01 \$12.00+\$2.00. US Postmaster send address changes to: Nature, PO Box 5055, Brentwood, TN 37024-5055. Published in Japan by Nature Japan K.K., Shin-Mitsuke Bldg. 36 Ichigaya Tamachi, Shinjuku-ku, Tokyo 162, Japan. © 2001 Nature Publishing Group.



# Visual transduction in *Drosophila*

Roger C. Hardie & Padinjat Raghu

Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK (e-mail: rch14@hermes.cam.ac.uk)

**The brain's capacity to analyse and interpret information is limited ultimately by the input it receives. This sets a premium on information capacity of sensory receptors, which can be maximized by optimizing sensitivity, speed and reliability of response. Nowhere is selection pressure for information capacity stronger than in the visual system, where speed and sensitivity can mean the difference between life and death. Phototransduction in flies represents the fastest G-protein-signalling cascade known. Analysis in *Drosophila* has revealed many of the underlying molecular strategies, leading to the discovery and characterization of signalling molecules of widespread importance.**

**P**hototransduction, the process by which light energy is converted into a photoreceptor's electrical response, has long been at the forefront of studies, not only of sensory transduction, but also cell signalling more generally. Pioneering studies in the 1970s and 80s unravelled the biochemical steps of excitation in vertebrate rods and, together with seminal studies of hormone-stimulated adenylate cyclase, led to the discovery and characterization of G-protein signalling<sup>1</sup>. These cascades, whereby heptahelical transmembrane receptors such as rhodopsin catalytically activate heterotrimeric G proteins, are widely found not only in many sensory receptors (see review in this issue by Firestein, pages 211–218), but also throughout the body, where they respond to all manner of chemical messengers, such as hormones, neurotransmitters, odorants and tastants.

## Photoreceptor performance

One hallmark of such cascades is their capacity for amplification. Early psychophysical experiments indicating that photoreceptors were capable of responding to single photons<sup>2</sup> were confirmed, first in invertebrates, and later in vertebrate rods, by electrophysiological recordings showing that quantized events (quantum bumps) could be recorded in response to absorption of single photons of light<sup>3,4</sup> (Fig. 1). Other functional attributes shared by vertebrate and invertebrate photoreceptors include low 'dark noise' (spontaneous thermal isomerizations of rhodopsin, which sets the ultimate limit on absolute sensitivity<sup>5</sup>); efficient mechanisms for response termination; the coding of intensity by graded potentials; and the ability to light adapt — that is, to reduce amplification as background intensity increases. But there are also differences that hint at a dichotomy in the underlying molecular machinery. First, vertebrate photoreceptors hyperpolarize, because the transduction channels close in response to light, whereas in most invertebrates the channels open, leading to depolarization. Second, in rods, the trade-off between amplification and response speed limits human temporal resolution to ~10 Hz under dim conditions. But fly photoreceptors possess the fastest known G-protein-signalling pathways, responding around 10 times more quickly than mammalian rods and ~100 times faster than toad rods recorded at similar temperatures (Fig. 1). Third, rods have only a limited ability to adapt, rapidly saturating as intensity increases; only the less sensitive cones can respond under daylight intensities. By contrast, despite their exquisite sensitivity to single

photons, fly photoreceptors successfully light adapt over the entire environmental range, up to ~10<sup>6</sup> effectively absorbed photons per second (Fig. 1)<sup>6–8</sup>.

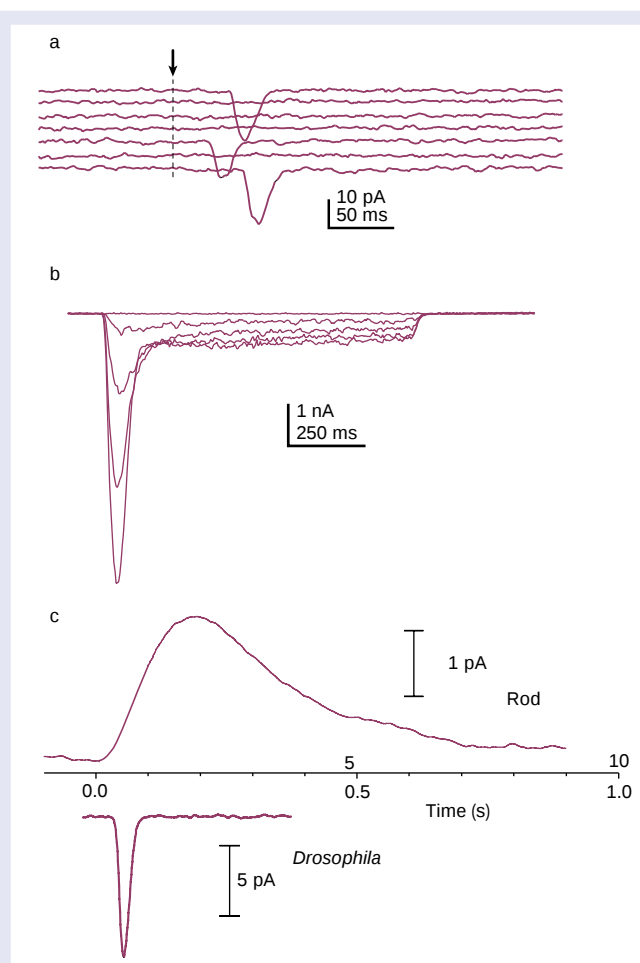
The phototransduction cascade in vertebrate rods is understood in unparalleled detail<sup>9,10</sup>, and widely cited as the textbook example of G-protein signalling, but the molecular strategies underlying invertebrate phototransduction are still being deduced. We focus here on recent studies in the fruitfly *Drosophila*, highlighting the similarities and differences with the well-established scheme in vertebrates. Key to our understanding is *Drosophila*'s unique genetic potential, which has been exploited to identify the elements of the cascade, while a powerful mix of molecular genetic and physiological analysis is providing insight into the molecular choreography by which these photoreceptors achieve their exceptional performance.

## Photoreceptor ultrastructure

Vertebrate and invertebrate photoreceptors both sequester their transduction machinery in specialized subcellular compartments (Fig. 2). Their structure is dictated in the first instance by the need to maximize the amount of light-absorbing membrane. Vertebrate rods achieve this with stacks of membranous discs internalized in the rod outer segment, which is separated from the rest of the cell by a short ciliary stalk. By contrast, invertebrate photoreceptors have tightly packed microvilli, which together form a cylindrical rhabdomere (from the Greek *rod*). Like its vertebrate counterpart, this acts as a light- or waveguide, trapping axially directed light, and at the same time contains most of the molecules of the transduction cascade. Converging studies indicate that individual microvilli, each only 60 nm in diameter, may be semiautonomous units of excitation and adaptation<sup>7,11,12</sup>. Together with their molecular organization, this miniaturization may be the key to understanding the amplification, rapid kinetics and adaptational capacity of these remarkable receptors.

## The visual cycle

Phototransduction begins with the absorption of light by rhodopsin, triggering the *11-cis* to *all-trans* photoisomerization of the chromophore (retinal or 2-dehydro-retinal in vertebrates, 3-hydroxy-retinal in flies<sup>13</sup>) and formation of the activated metarhodopsin state. In vertebrates, *all-trans* retinal subsequently dissociates and must be re-isomerized through a lengthy and time-consuming enzymatic pathway that dictates the time course of dark adaptation following bleaching illumination (~30 min for rods). Invertebrate



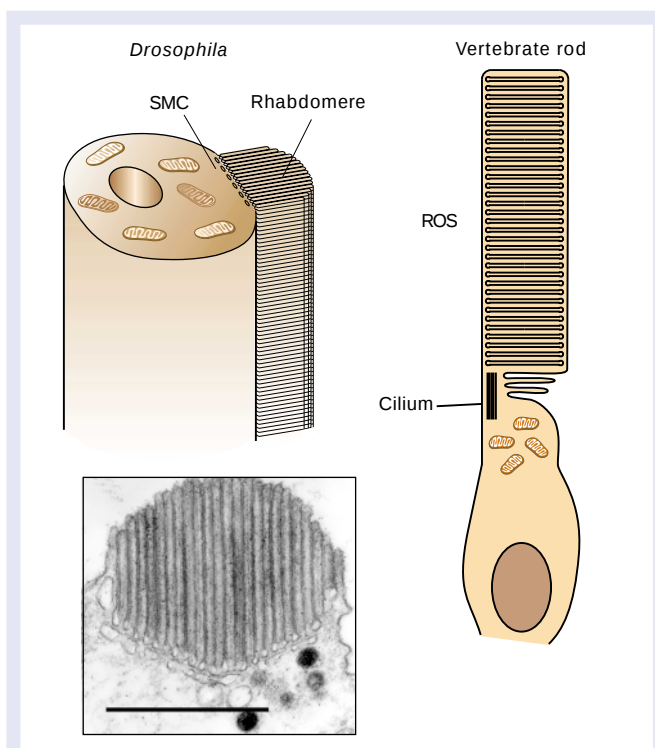
**Figure 1** Photoreceptor responses. The physiological analysis of phototransduction in *Drosophila* was revolutionized 10 years ago by the development of a preparation allowing whole-cell patch-clamp analysis of the light-sensitive current<sup>86,87</sup>.

**a.** Voltage-clamped responses to brief, dim light flashes containing on average less than a single effective photon elicit discrete quantized inward currents, known as quantum bumps. **b.** With longer-duration (1 s) flashes of increasing intensity, the bumps fuse to form a noisy inward current. At higher intensities, light adaptation is manifested as a rapid transition from a peak, which can reach values in excess of 20 nA, to a steady-state plateau of ~500 pA. **c.** Comparison of quantum-bump kinetics in toad rod outer segments (data from ref. 88, courtesy T. D. Lamb) and *Drosophila* recorded at 21 °C (note the different timescales).

metarhodopsin is usually thermostable, and can be directly re-isomerized back to rhodopsin by absorption of longer wavelength light. Fly eyes are red because the retinal screening pigments are transparent to long wavelengths: this represents a particularly economic strategy, allowing metarhodopsin to be constantly reconverted back to rhodopsin by ambient light filtering diffusely through the eye tissue<sup>14</sup>.

### Invertebrates use the phosphoinositide pathway

In vertebrate rods, the heterotrimeric G protein transducin activates a phosphodiesterase (PDE) resulting in hydrolysis of guanosine 3',5'-cyclic monophosphate (cGMP) and closure of the transduction channels (Box 1). In *Drosophila*, as in most invertebrates, rhodopsin activates a distinct G-protein isoform,  $G_q$ , which activates, instead of PDE, a phospholipase C isoform (PLC $\beta$ 4, encoded by the *norpA* gene). This leads to opening of two classes of  $Ca^{2+}$ -permeable light-sensitive channels: transient receptor potential (TRP) and TRP-like (TRPL) channels<sup>15–17</sup>. PLC is well known as the effector enzyme of the phosphoinositide pathway (Box 1), generating soluble



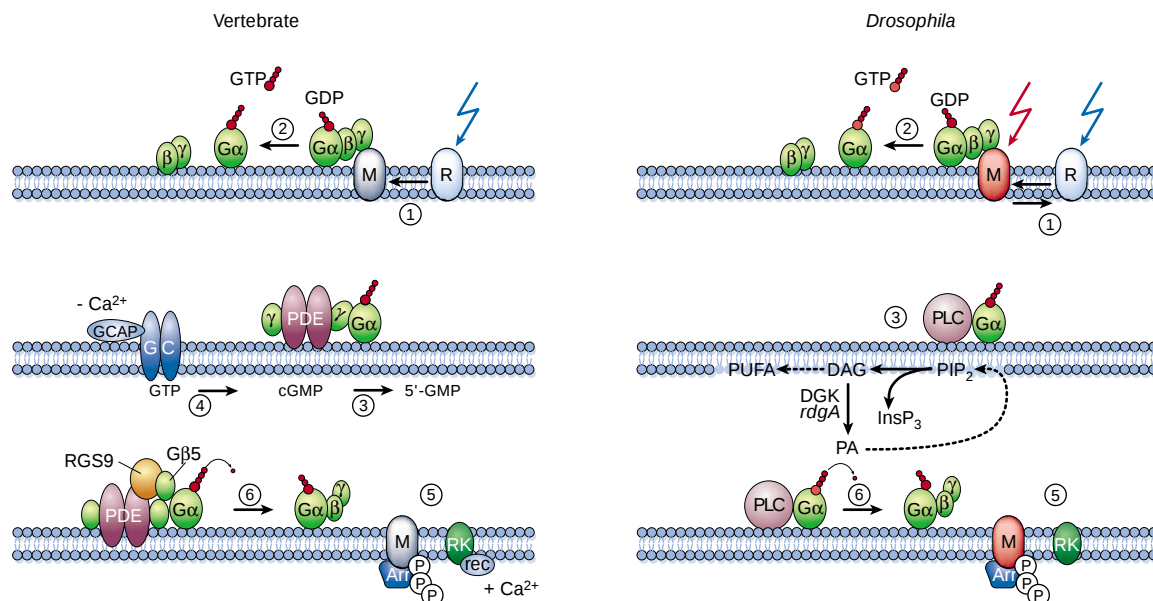
**Figure 2** Photoreceptor structure. In *Drosophila*, as in most invertebrate photoreceptors, the photoreceptive membrane is organized into tightly packed, tubular microvilli, each 1–2  $\mu$ m long and ~60 nm in diameter, together forming a 100- $\mu$ m-long rhabdomere. At the base of the microvilli a system of submicrovillar cisternae (SMC) have often been presumed to represent smooth endoplasmic reticulum  $Ca^{2+}$  stores endowed with  $Ins(1,4,5)P_3$  receptors. However, the SMC may have a more important role in phosphoinositide turnover. Vertebrate rod outer segments (ROS) contain stacks of membranous discs (~1,000) and are connected to the cell body by a narrow cilium. In both cases the overall structure serves to maximize absorption of light by forming a cylindrical light-guiding structure with a high density of rhodopsin-containing membrane. Inset shows electron micrograph of one rhabdomere (courtesy of A. Polyanovsky; scale bar, 1  $\mu$ m).

inositol-1,4,5-trisphosphate ( $Ins(1,4,5)P_3$ ) and membrane-bound diacylglycerol (DAG) from hydrolysis of the minor membrane phospholipid phosphatidylinositol-4,5-bisphosphate ( $PtdIns(4,5)P_2$ ). With numerous mutants in over 20 cloned genes, *Drosophila* phototransduction represents the best genetic model of this ubiquitous ' $Ca^{2+}$ -signalling' pathway among higher eukaryotes.

### Evidence for a lipid messenger of excitation

The essential role of PLC is undisputed, but the downstream events leading to gating of the transduction channels are controversial — a situation that will be familiar to those attempting to understand the analogous process of PLC-regulated  $Ca^{2+}$  influx in many other cells<sup>18</sup>. Release of  $Ca^{2+}$  from  $Ins(1,4,5)P_3$ -sensitive stores is believed to be an essential step in photoreceptors of some invertebrates, such as *Limulus*<sup>19</sup>, but apparently not in *Drosophila* as phototransduction is unaffected in mutants of the only  $Ins(1,4,5)P_3$ -receptor gene in the genome<sup>20,21</sup>. This has redirected attention to membrane-delimited consequences of PLC activity, namely, generation of DAG or the reduction in  $PtdIns(4,5)P_2$  levels. DAG is best known as an activator of protein kinase C (PKC), but mutants in an eye-specific PKC have defects only in response inactivation and adaptation, leaving excitation unaffected<sup>22,23</sup>. DAG also is a potential substrate for DAG lipase, leading to release of polyunsaturated fatty acids (PUFAs) such as arachidonic acid. Application of PUFAs activates TRP and TRPL channels *in situ* and recombinant TRPL channels can be activated by both PUFAs<sup>24</sup> and by DAG itself<sup>25</sup>.

## Box 1

Phototransduction in cascades in vertebrate rods and *Drosophila*


The encircled numbers in the figure above refer to the following steps in the transduction cascades. (1) Photoisomerization. Rhodopsin (R) is photoisomerized to metarhodopsin (M). *Drosophila* M is thermostable, and can be reconverted to R by long-wavelength light; vertebrate M releases the bound chromophore (*all-trans* retinal). (2) GTP/GDP exchange. M catalyses exchange of GDP for GTP on the heterotrimeric G protein (transducin in rods,  $G_q$  in *Drosophila*); active GTP-bound  $\alpha$ -subunit dissociates. (3) Activation.  $G_\alpha$  binds to and activates the effector enzyme (PDE in rods, PLC in *Drosophila*). In vertebrates, PDE hydrolyses cGMP to 5'-GMP, leading to closure of CNG channels. In *Drosophila*, PLC hydrolyses PtdIns(4,5) $P_2$  (PIP<sub>2</sub>) to DAG and Ins(1,4,5) $P_3$ . DAG is also a potential substrate for DAG lipase, leading to the release of PUFAs. Two

classes of channel (TRP and TRPL) are activated by an unknown mechanism. (4) Substrate resynthesis. In vertebrates, cGMP is resynthesized by guanylate cyclase (GC) and GC-activating protein (GCAP), which is inhibited by  $Ca^{2+}$ . In *Drosophila*, DAG is converted to phosphatidic acid (PA) by DAG kinase (DGK). PA is converted to PtdIns(4,5) $P_2$  by a multienzymatic pathway. (5) Metarhodopsin inactivation. M is phosphorylated by rhodopsin kinase (RK) and capped by arrestin. RK is inhibited by recoverin in presence of  $Ca^{2+}$  (vertebrates only). (6) Inactivation of G protein and effector. Effector enzyme and  $G_\alpha$  are inactivated by the GTPase activity of the G protein, leading to reassociation with  $G\beta\gamma$ . Accelerated by the GAP activity of RGS9/ $G\beta 5$  and PDE (rods) and PLC (*Drosophila*).

Independent evidence that DAG may be important in excitation comes from a blind retinal degeneration mutant, *rdgA*. The *rdgA* gene encodes DAG kinase (DGK)<sup>26</sup>, which controls DAG levels by converting it to phosphatidic acid. TRP channels are constitutively active in *rdgA* mutants and the resulting  $Ca^{2+}$  influx may trigger the degeneration, as this is rescued in *rdgA;trp* double mutants<sup>27</sup>. The response to light is restored in these double mutants, but deactivates abnormally slowly, which indicates that DGK is required for response termination. The deactivation defect and constitutive activity are consistent with a role for DAG in excitation. However, DGK is also involved in synthesis of PtdIns(4,5) $P_2$  (Box 1), so that PtdIns(4,5) $P_2$  levels and the kinetics of its recycling may also be impaired in *rdgA* mutants. PtdIns(4,5) $P_2$  regulates the activity of a number of ion channels, including the Kir family of inward rectifiers<sup>28</sup> and the TRP-related vanilloid receptor<sup>29</sup>. Recombinant TRPL-channel activity was recently reported to be suppressed by application of PtdIns(4,5) $P_2$  (ref. 25), whereas PtdIns(4,5) $P_2$  depletion was shown to be correlated with activation of the photoreceptor TRP channels *in situ*<sup>30</sup>. These findings indicate that reduction in PtdIns(4,5) $P_2$  should be considered as a potential mechanism of channel gating.

Final resolution of the mechanism of excitation in *Drosophila* is likely to require identification and characterization of ligand-binding sites on the channel molecules, analysis of light-induced lipid metabolism, and identification and mutant analysis of any further genes products required for activation. For example, a

mutation of a completely novel protein (INAF) has recently been shown to mimic aspects of the *trp* phenotype, suggesting it may be required for TRP activation<sup>31</sup>.

#### Possible phosphoinositide signalling in vertebrate rods

Surprisingly, vertebrate photoreceptors also express a PLC $\beta 4$  isoform, which is more closely related to *Drosophila* *norpA* than it is to other vertebrate PLC isoforms, perhaps indicative of a common ancestral photoreceptor in the distant evolutionary past<sup>32</sup>. Its function is unknown, but there are reports of light-induced phosphoinositide metabolism in rods, and PtdIns(4,5) $P_2$  has recently been shown to inhibit the rod cyclic nucleotide-gated (CNG) channels and stimulate PDE<sup>33</sup>.

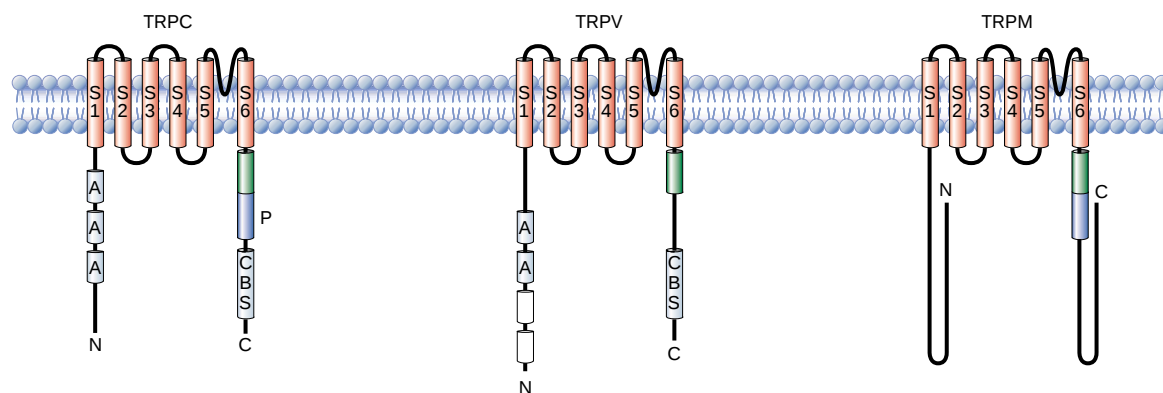
#### Response termination

In any transduction cascade it is essential that each component be efficiently inactivated. Failure to do so, for example in specific mutants, results in long-lasting responses, compromising temporal resolution<sup>34</sup>. A common theme in G-protein signalling is that the receptor is inactivated by binding to arrestin (Box 1). In vertebrate rods, metarhodopsin must also first be multiply phosphorylated by rhodopsin kinase, and elimination of the responsible serine residues has shown these are required for reliable and rapid deactivation<sup>35,36</sup>. The carboxy terminal of *Drosophila* metarhodopsin is similarly phosphorylated, but the function is obscure as mutants lacking the phosphorylation sites show normal response kinetics<sup>37</sup>; arrestin



## Box 2

## TRP channels



Most closely related to *Drosophila trp*. Seven mammalian isoforms (TRPC1-7). Activated downstream of PLC. Short N termini with three *ank* repeats.

Five mammalian isoforms, including VR1 (TRPV1) involved in thermal nociception (see review in this issue by Julius and Basbaum). Often gated directly by chemical or physical stimuli. Between two and four *ank* repeats; lack of proline-rich region after S6.

Eight mammalian isoforms, including melastatin (TRPM1). Long N termini lack *ank* repeats. TRPM2 (ref. 95) and TRPM7 (TRP-PLIK)<sup>96</sup> harbour enzymatic domains in their C termini.

TRP-related ion channels are one of the most recently discovered ion-channel families with ~20 mammalian isoforms. Many are involved in sensory transduction not only in vision, but also in olfaction, pain, and mechano- and osmosensation (see reviews in this issue by Gillespie and Walker, Julius and Basbaum, and Firestein). More generally, they are responsible for a wide range of  $\text{Ca}^{2+}$ - and cation-influx pathways. TRP-related ion channels are divided into at least three subfamilies — TRPC (formerly STRP), TRPV (or OTRP) and TRPM (or LTRP)<sup>52,53</sup>; see figure above. The basic topology of TRPs, usually with six transmembrane domains (S1–6), is typical of the superfamily of voltage-gated and CNG channels. By analogy with

voltage-gated  $\text{K}^+$  channels and CNG channels, TRPs are presumed to form subunits of tetrameric channels. The most conserved features of the TRP family include the N terminus, which contains multiple ankyrin repeats (absent in TRPMs), and a ~50-residue 'TRP domain' followed by a proline-rich domain (P) just after the last transmembrane helix. Several TRPs, including dTRP<sup>90</sup>, dTRPL<sup>92</sup> and TRPC3<sup>93</sup>, and at least one TRPV (CaT-L)<sup>94</sup> contain one or more CaM-binding sites (CBS) in the C terminus, which have been implicated in  $\text{Ca}^{2+}$ -dependent inactivation. In general, however, the C-terminal regions show little sequence similarity, consistent with the functional diversity of the family.

binding is, however, essential to quench metarhodopsin activity<sup>38,39</sup>. Activity of G protein and effector enzyme is terminated by the GTPase activity of the G protein. But the intrinsic GTPase activity is too slow to account for the rapid response termination, and it is now clear that binding to the effector enzyme itself (PDE in rods, PLC in flies) is required to accelerate GTP hydrolysis<sup>40</sup>. This requirement has an elegant logic in that the G protein will not be inactivated until it has first encountered and activated its downstream effector. In vertebrate rods, additional binding of a specific GTPase-activating protein (GAP), regulator of G-protein signalling 9 (RGS9; complexed with Gβ5), is also required<sup>41</sup>. Whether RGS proteins have similar roles in *Drosophila* photoreceptors is unknown. The final inactivation step involves the channels: in vertebrates these must be re-opened by synthesis of cGMP by particulate guanylate cyclase. Control of this enzyme by  $\text{Ca}^{2+}$ -dependent feedback acting through a small  $\text{Ca}^{2+}$ -binding protein, guanylate cyclase activating protein (GCAP), is one of the main mechanisms of light adaptation<sup>10</sup>.

### Transduction channels

The transduction channels have a central role in both vertebrates and invertebrates. In addition to mediating the electrical response, they are highly permeable to  $\text{Ca}^{2+}$ , which is a key mediator of response termination and adaptation. When first cloned, these channels were found to define new classes of ion channel and their detailed analysis have provided essential clues to the mechanism of transduction.

### Vertebrate CNG channels

The discovery that the transduction channels of vertebrate rods were gated by cGMP in inside-out patches was the clinching argument in the long-running debate over the identity of the second messenger in

vertebrate phototransduction<sup>42</sup>. When cloned, the  $\alpha$ -subunit of the rod CNG channel defined a new family, with six mammalian isoforms, within the superfamily of voltage-gated ion channels possessing six transmembrane domains and a cGMP-binding site<sup>43</sup>. Related CNG channels are found in a variety of neuronal and non-neuronal tissue, including distinct isoforms in cones and the transduction channels in olfactory receptors (see review in this issue by Firestein, pages 211–218). The native rod channel is a heteromultimer with a  $\beta$ -subunit, notable for a glutamate acid-rich protein (GARP) sequence in the cytoplasmic tail and a calmodulin (CaM)-binding domain responsible for modulating the affinity of the channel for cGMP during light adaptation<sup>44</sup>.

An unusual property of the photoreceptor CNG channels is a tiny single-channel conductance (~100 fS), due to a voltage-dependent divalent ion block. This may improve signal-to-noise ratio by allowing a much larger number of channels (~10,000) to be simultaneously active than would otherwise be possible<sup>45</sup>. But perhaps the most significant functional property of the CNG channels is their high  $\text{Ca}^{2+}$  permeability.  $\text{Ca}^{2+}$  levels in rods and cones are controlled dynamically by the balance between  $\text{Ca}^{2+}$  influx through the CNG channels and  $\text{Ca}^{2+}$  extrusion by the  $\text{Na}^+/\text{Ca}^{2+}/\text{K}^+$  exchanger. As the channels close in response to light,  $\text{Ca}^{2+}$  continues to be extruded and the resulting reduction is the essential feedback signal facilitating response termination and mediating light adaptation<sup>10,46</sup> (Box 1).

### *Drosophila* TRP channels

In contrast to vertebrate rods, which contain only one functional class of transduction channel, *Drosophila* photoreceptors express at least two distinct channels encoded by up to three genes. The *trp* gene is required for the main component of the light-sensitive

conductance, which, *in vivo*, has a high  $\text{Ca}^{2+}$  selectivity ( $P_{\text{Ca}}:P_{\text{Na}} > 100:1$ )<sup>47–49</sup>. A second conductance is mediated by a non-selective cation channel encoded by a homologous gene, *trpl*<sup>17,50</sup>, possibly in heteromultimeric combination with a third recently discovered homologue, *trp-γ*<sup>51</sup>. The predicted sequences of *trp*, *trpl* and *trp-γ* have six transmembrane  $\alpha$ -helices, representing putative subunits of multimeric channels, and define a new class of channel, again within the same superfamily of voltage-gated and CNG channels. The extended TRP family (Box 2) includes an eclectic collection of ion channels. Those most closely related to *Drosophila* TRP also seem to be activated by PLC pathways; others include several other candidate sensory-transduction channels (see refs 52, 53, and reviews in this issue by Julius and Basbaum, pages 203–210, Gillespie and Walker, pages 194–202, and Firestein, pages 211–218).

Although only distantly related, some intriguing functional similarities between *Drosophila* TRP and rod CNG channels may reflect convergent roles in phototransduction. TRP channels are highly  $\text{Ca}^{2+}$  permeable, and  $\text{Ca}^{2+}$  influx via TRP channels mediates amplification, rapid response termination and light adaptation through multiple feedback targets<sup>34,54</sup>. The  $\text{Ca}^{2+}$  influx has also been implicated recently in regulating  $\text{PtdIns}(4,5)\text{P}_2$  metabolism by inhibiting PLC and facilitating  $\text{PtdIns}(4,5)\text{P}_2$  recycling<sup>30</sup>. As in CNG channels, TRP and TRPL harbour one and two  $\text{CaM}$ -binding sites, respectively; at least in TRPL, these seem to be involved in  $\text{Ca}^{2+}$ -dependent inactivation of the channel<sup>55</sup>. Like CNG channels, TRP is also subject to a voltage-dependent divalent ion block, but with a subtly different functional outcome: the block intensifies as the cell depolarizes over the physiological range of voltages (–70 to 0 mV) and would seem to represent an elegant and economic mechanism for reducing gain during light adaptation<sup>56</sup>.

### Diffusion versus signalling complexes

Amplification in vertebrate rods is believed to rely upon sequential, stochastic diffusional encounters. Rhodopsin first activates several hundred transducin molecules during a random walk in the disc membrane; activated transducin  $\alpha$ -subunits then immediately start binding to and activating PDE molecules, again by diffusional encounters. The catalytic power of PDE is among the highest known of any enzyme, and is effectively limited only by the diffusional access of cGMP<sup>57</sup>. The high density of rhodopsin in the disc membrane (essential to maximize absorption of light) actually hinders its own diffusion. Correspondingly, when rhodopsin concentration is halved in hemizygote  $\text{Rh}^{+/+}$ -knockout mice, the kinetics of both excitation and deactivation are accelerated about twofold, in agreement with the predicted enhanced mobility of the more sparsely distributed rhodopsin<sup>58</sup>.

**Table 1 Approximate stoichiometry of components of signalling cascades**

|                   | R | G protein | Effector (PDE/PLC) | Channel (CNG/TRP) | Channels per bump* | Max current (pA) |
|-------------------|---|-----------|--------------------|-------------------|--------------------|------------------|
| Vertebrate rod    | 1 | 1/10      | 1/100              | 1/1000†           | 250                | 10–50            |
| <i>Drosophila</i> | 1 | 1/10      | 1/10               | 1/40‡             | 15                 | >20,000§         |

Approximate stoichiometry of the main components of the cascades expressed as mole fraction of rhodopsin (R;  $\sim 10^5$  molecules per disc, and  $10^3$  molecules per microvillus)<sup>63,64</sup>. *Drosophila*'s strategy requires each of the 30,000 microvilli to have a high density of G protein and PLC and sufficient channels to generate a quantum bump. The maximum light-induced current is about 1,000-times greater than in rods. Although the steady-state current during light adaptation is a more modest 500 pA, this is still ten-times greater than the maximum dark current found in vertebrates. Because ion fluxes are usually the largest drain on a nerve cell's energy budget<sup>65</sup>, *Drosophila*'s strategy is likely to be expensive.

\*Mean bump amplitude divided by single-channel current.

†A maximum of only  $\sim 2\%$  are open at one time.

‡Assumes four TRP proteins per channel.

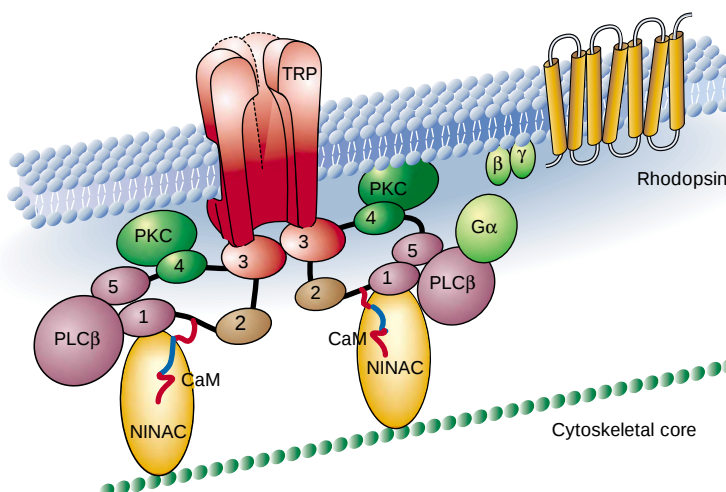
§Minimum estimate; maximum currents are too large to be voltage-clamped accurately.

The diffusional model has been an influential concept of intracellular signalling. But just how general is the concept of randomly interacting proteins as a principle of signal transduction? Increasing evidence indicates that receptors, enzymes and channels may, instead, often be organized into multimolecular signalling complexes<sup>59</sup>. By assembling elements in specific subcellular localizations, these could promote speed, specificity and reproducibility of response. Such complexes are often organized around 'scaffolding' proteins containing one or more 'PDZ' domains. Named after PSD-95 (postsynaptic density protein), *Drosophila* discs large (*dlg*), and the tight-junction protein ZO-1, PDZ domains are protein modules of about 90 amino acids that bind to a variety of target proteins by means of specific target sequences, such as an S/T-X-V/I motif in the final three amino acids of the C terminus<sup>60</sup>.

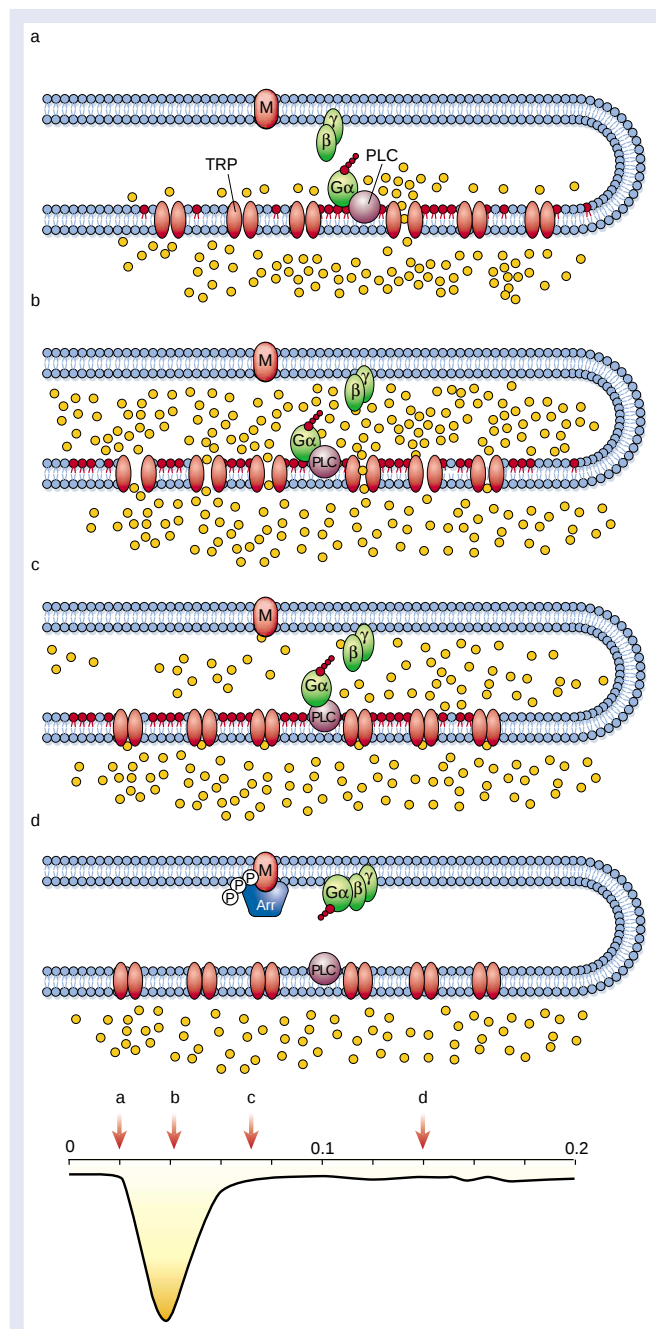
### The INAD complex

Several key elements of the *Drosophila* phototransduction cascade are now thought to be assembled into a multimolecular complex by a scaffolding protein, INAD, with a total of five PDZ domains (Fig. 3). Analysis of this complex, dubbed 'transducisome'<sup>61</sup> or 'signalplex'<sup>62</sup>, provides the best example of how a PDZ-domain protein organizes a complex signalling cascade. Three key components of the cascade — the TRP channel, PLC and an eye-specific PKC required for response termination and light adaptation — form the core of a macromolecular complex by binding to individual PDZ domains on INAD<sup>61,63,64</sup>. INAD is required for the correct microvillar localization of all these components. PKC and PLC require INAD for correct targeting and seem to be pre-assembled with it before being transported to the rhabdomere<sup>65</sup>. By contrast, TRP and INAD are initially correctly targeted to the microvilli in each other's absence, but become delocalized with age, indicating a reciprocal requirement for long-term retention<sup>64,65</sup>. As the

**Figure 3** The INAD signalling complex. The INAD protein contains five PDZ domains (1–5) joined by short linker regions. Each PDZ domain associates preferentially with different targets<sup>61,63,66,89,90</sup>. The precise composition of the native complex is uncertain, as some PDZ domains are reported to bind at least two different targets. There are also several possibilities for multimerization, for example by homophilic interactions between INAD (PDZ domains 3 and 4)<sup>66</sup>, or by involvement of up to four TRP subunits and linkage of two INAD molecules via PLC $\beta$ , which binds both PDZ1 and PDZ5<sup>68</sup>. In the model shown here, two INAD complexes are shown bound both by two TRP subunits and by their PDZ3 domains, whereas PLC $\beta$  is shown linking PDZ1 and PDZ5 of the same INAD molecule. Calmodulin (CaM) binds to the linker region between PDZ1 and PDZ2<sup>66</sup>. The  $G_q$   $\alpha$ -subunit acts as a diffusible shuttle between activated rhodopsin and the complex.



only transmembrane protein in the complex, TRP may be required for anchoring to the membrane. Importantly, INAD probably also determines the stoichiometry of the transduction machinery, as quantitative western-blot analysis indicates that PKC, PLC, INAD and TRP are expressed in approximately equal numbers<sup>63</sup> (Table 1).



**Figure 4** A model of bump generation in *Drosophila*. **a**, Approximately 20 ms after absorption of a photon, metarhodopsin (M) has activated at least one G protein, which in turn has activated PLC, generating a membrane-delimited second messenger (this could be DAG, PUFA or reduction in PtdIns(4,5)P<sub>2</sub>) indicated in red. Threshold for activation of one channel is reached, whereas channels further from PLC 'see' only subthreshold amounts. **b**, But Ca<sup>2+</sup> influx rapidly raises [Ca<sup>2+</sup>] in the microvillus, and sensitizes the remaining channels, possibly by increasing affinity for the putative second messenger; this positive feedback generates the abrupt rising phase of the bump. **c**, Ca<sup>2+</sup> floods the whole microvillus (>200 μM), leading to rapid inactivation and a refractory period. **d**, Ca<sup>2+</sup> is returned to resting levels (~150 nM) within ~100 ms. M, Gα and PLC are deactivated and PtdIns(4,5)P<sub>2</sub> resynthesized during this refractory period. See text for further details.

Four further elements, namely the second transduction channel TRPL, CaM, rhodopsin and the NINAC protein (a myosin III potentially linking the complex to the cytoskeleton), also bind to INAD, although none of these proteins seem to require interaction with INAD for microvillar localization<sup>66,67</sup>. Finally, INAD multimerizes *in vitro* through homophilic PDZ interactions; together with the TRP protein, which also forms multimers, this might give rise to an extended web of INAD complexes ('signalplex')<sup>66</sup>.

What is the significance of this complex molecular architecture? INAD seems essential for correct microvillar localization and probably the stoichiometric relation of transduction molecules. But it is unclear whether this alone is sufficient for normal transduction or whether specific protein–protein interactions within the complex are also crucial. With respect to TRP, presence of the channels alone may be sufficient, as young flies expressing a mutant TRP construct incapable of binding to INAD show no obvious response defects, developing a phenotype only when the TRP protein becomes delocalized with age<sup>64</sup>.

There are some indications that INAD's role might be more direct and possibly more dynamic than simply being required for appropriate subcellular localization. First, INAD has more reported protein partners than PDZ domains<sup>67</sup>, raising the possibility of dynamic competition for binding sites. Second, PLC is reported to bind to two different PDZ domains on INAD by distinct binding sites, one of which overlaps with the G-protein-binding domain, raising the possibility that PLC activation may be regulated by the PDZ interaction<sup>68</sup>. Third, both TRP and INAD can be phosphorylated by PKC, another integral member of the complex. Clearly, direct juxtaposition of kinase and substrate is likely to promote rapid and efficient phosphorylation, and also raises the possibility that PDZ–target interactions could be regulated by phosphorylation<sup>69,70</sup>.

#### Signalling complexes in vertebrate photoreceptors

Although diffusional encounters probably underlie the rate-limiting steps of activation and inactivation in vertebrate phototransduction<sup>58</sup>, recent studies have indicated that multimolecular complexes may also be important. First, the active state of PDE, itself bound in a complex with G<sub>t</sub>α, is deactivated by binding to another complex comprising RGS9 and Gβ5, which together with the PDE γ-subunit confer the necessary GAP activity for GTP hydrolysis<sup>71</sup>. Second, the rod CNG channels seem to be part of a complex including the Na<sup>+</sup>/Ca<sup>2+</sup>/K<sup>+</sup> exchanger, which binds to the CNG α-subunit<sup>72</sup>, and the disc-rim protein, peripherin, which associates with the GARP domain of the CNG β-subunit (R. S. Molday, personal communication). Soluble GARP proteins may also be part of the complex, and have been reported to associate with the light-activated state of PDE, the interaction causing up to fivefold inhibition of PDE<sup>73</sup>. Possible roles for this complex, which is found only in rods but not in cones, include spatial localization of Ca<sup>2+</sup> transients, anchoring and spacing of the discs, and downregulation of cGMP turnover during daylight when rod function is saturated.

#### A molecular strategy for transduction in *Drosophila*

How does the molecular and cellular organization of the *Drosophila* photoreceptor achieve the impressive combination of amplification, rapid response kinetics and adaptational capacity, which sets it apart from the vertebrate rod? Analysis of quantum bumps suggests a fundamentally different strategy. In rods, bumps arise with almost negligible latency but a relatively slow rise time, which can be accurately modelled by a diffusion-limited amplification scheme<sup>74</sup>. By contrast, in *Drosophila* there is a finite, although variable latency, following which the bump rises and falls abruptly (Fig. 1). Several lines of evidence indicate that this reflects a mechanism more akin to an action potential, involving a threshold, positive and negative feedback by Ca<sup>2+</sup>, and a refractory period<sup>75</sup>. A key finding is that in mutants where PLC levels are reduced, bump latency can be increased markedly (to ~1 min), whereas bump amplitude is



unaffected<sup>76,77</sup>. This implies that all amplification is mediated downstream of PLC, that latency is determined by events up to and including PLC, and that a single G protein and PLC may be sufficient to generate a bump. Although not absolutely essential for the following model (Fig. 4), we also assume that excitation is restricted to one microvillus.

### Latency

Bump latency, which can be as short as 20 ms, represents the time taken to activate PLC and generate a sufficient quantity of second messenger to activate the first channel. In contrast to rods, rhodopsin in microvillar photoreceptors is essentially non-diffusible. Because INAD complexes are also likely to have restricted mobility, it seems that  $G_q\alpha$ , which forms transient interactions with the complex by binding to PLC<sup>78</sup>, must act as a diffusible shuttle between rhodopsin and the complexes<sup>78,79</sup>. Factors that promote a short latency probably include: (1) the need to activate at most a few G proteins, which, together with the high concentration of rhodopsin and PLC in the microvillus, minimizes diffusional delays; (2) the close proximity of channels to PLC; and (3) a high enzymatic activity of PLC<sup>30</sup>.

### $Ca^{2+}$ mediates positive feedback

Threshold is reached when the first channels start to open, resulting in  $Ca^{2+}$  influx into the microvillus. Because of its restricted volume,  $Ca^{2+}$  rises almost instantaneously throughout the microvillus, initiating rapid positive and negative feedback. Within 10–20 ms of the first channel opening, activation is maximal with ~15 TRP channels open at the peak of the bump<sup>12</sup>. This may represent activation of most of the channels in the microvillus — in essence a regenerative ‘all-or-none’ event. Positive feedback is a unique feature of invertebrate phototransduction, and is evident from the profound (tenfold) reduction in bump amplitude in the absence of external  $Ca^{2+}$  (ref. 12).

### Termination and refractory period

The microvillus is now flooded with  $Ca^{2+}$  in excess of 200  $\mu$ M (ref. 80). This results in rapid, complete  $Ca^{2+}$ -dependent inactivation of the channels. PKC, which is activated by  $Ca^{2+}$  and DAG, has been implicated in this behaviour, as quantum bumps in mutants lacking PKC show severe defects in termination<sup>23</sup>. Perhaps the close proximity of TRP and PKC within the INAD complex allows rapid phosphorylation of the channel, making it amenable to  $Ca^{2+}$ -dependent inactivation. It seems likely that the exceptionally high  $Ca^{2+}$  levels also render the microvillus temporarily refractory to further stimulation, until after  $Ca^{2+}$  is removed — within ~100 ms in the dark and more quickly when light adapted<sup>80</sup>. A particular advantage of such a refractory period may be to allow the ‘off’ kinetics to be determined by the rapid  $Ca^{2+}$ -dependent inactivation. Potentially slower biochemical events (such as arrestin binding, GTPase activity and PtdIns(4,5)P<sub>2</sub> resynthesis) need not then be rate limiting as long as they are completed within the refractory period. Support for a refractory period comes from paired flash experiments<sup>11</sup> and observation of quantum bumps in mutants where rhodopsin fails to inactivate. In vertebrates, bumps in such mutants simply fail to terminate and last for several seconds<sup>35</sup>. But in *Drosophila* mutants, a single photon absorption induces a train of normal shaped bumps separated by ~100–200 ms (ref. 55), presumably representing the refractory period.

### From single quanta to sunlight

How can *Drosophila* photoreceptors respond over the wide range of environmental intensities? The answer may lie in the localization of excitation to single microvilli and the short refractory period. Even during light adaptation, noise analysis indicates that the response can be accounted for largely by the linear summation of bumps; these become smaller and faster with increasing background adaptation, owing to negative feedback from raised steady-state  $Ca^{2+}$  levels<sup>6,8,81</sup>. With ~30,000 microvilli, each of which can be recycled at least every

100 ms, this should allow the photoreceptors to handle photon fluxes in excess of 300,000 s<sup>-1</sup>, as observed experimentally<sup>7,8</sup>. If this is not enough, flies possess one additional adaptation. In every photoreceptor there are tiny pigment granules, ~0.2  $\mu$ m in diameter; in response to the light-induced rise in  $Ca^{2+}$  they migrate towards the base of the rhabdomere<sup>82</sup> where they attenuate the light flux by up to two orders of magnitude, preventing saturation under even the brightest daylight intensities<sup>7,83</sup>.

### Conclusions and perspectives

An important challenge of the postgenomic era is to understand how defined molecular components interact to generate a physiological output. The precision with which stimuli can be controlled, and responses measured, has always made photoreceptors preferred systems for pursuit of this goal. Indeed, excitation kinetics in vertebrate rods were modelled successfully from the properties of identified molecular components some 10 years ago<sup>84</sup>. But before such a quantitative account can be attempted for phototransduction in *Drosophila*, outstanding questions still need to be addressed. The most urgent is the mechanism of activation of the transduction channels. The final response of the cell is orchestrated by rapid and complex  $Ca^{2+}$ -dependent feedback within the context of the supramolecular organization coordinated by INAD molecules. Even though this is the best example of such an organized signalling cascade, the detailed architecture and functional significance of these complexes, as well as the molecular mechanisms of  $Ca^{2+}$ -feedback regulation, still remain controversial.

How many genes are required to mediate phototransduction? Most of the genes known to be important in *Drosophila* were isolated by screening for viable mutants, so that any lethal mutations may have been overlooked. In addition, it has been estimated that two-thirds of all genes in *Drosophila* when mutated will have no phenotype<sup>85</sup>. With the recent completion of the *Drosophila* genome project and the development of a new generation of functional genomic tools, the prospect of identifying these remaining molecules is greatly improved. Together with the unique experimental accessibility of the retina, this should guarantee many exciting findings in the coming years, not only in phototransduction, but also in the cell biology of calcium signalling, lipid messengers and neuronal degeneration. □

- Hille, B. G protein-coupled mechanisms and nervous signaling. *Neuron* **9**, 187–195 (1992).
- Hecht, S., Shlaer, S. & Pirenne, M. Energy quanta and vision. *J. Gen. Physiol.* **25**, 819–840 (1942).
- Yeandle, S. & Spiegel, J. B. Light-evoked and spontaneous discrete waves in the ventral nerve photoreceptor of *Limulus*. *J. Gen. Physiol.* **61**, 552–571 (1973).
- Baylor, D. A., Lamb, T. D. & Yau, K.-W. Responses of retinal rods to single photons. *J. Physiol.* **288**, 613–634 (1979).
- Aho, A. C., Donner, K., Hyden, C., Larsen, L. O. & Reuter, T. Low retinal noise in animals with low body-temperature allows high visual sensitivity. *Nature* **334**, 348–350 (1988).
- Wu, C.-F. & Pak, W. L. Light-induced voltage noise in the photoreceptor of *Drosophila melanogaster*. *J. Gen. Physiol.* **71**, 249–268 (1978).
- Howard, J., Blakeslee, B. & Laughlin, S. B. The intracellular pupil mechanism and the maintenance of photoreceptor signal:noise ratios in the blowfly *Lucilia cuprina*. *Proc. R. Soc. Lond. B* **231**, 415–435 (1987).
- Juusola, M. & Hardie, R. C. Light Adaptation in *Drosophila* photoreceptors. I. Response dynamics and signaling efficiency at 25 °C. *J. Gen. Physiol.* **117**, 3–25 (2001).
- Lamb, T. D. Gain and kinetics of activation in the G-protein cascade of phototransduction. *Proc. Natl Acad. Sci. USA* **93**, 566–570 (1996).
- Koutalos, Y. & Yau, K. W. Regulation of sensitivity in vertebrate rod photoreceptors by calcium. *Trends Neurosci.* **19**, 73–81 (1996).
- Hochstrate, P. & Hamdorf, H. Microvillar components of light adaptation in blowflies. *J. Gen. Physiol.* **95**, 891–910 (1990).
- Henderson, S. R., Reuss, H. & Hardie, R. C. Single photon responses in *Drosophila* photoreceptors and their regulation by  $Ca^{2+}$ . *J. Physiol.* **524**, 179–194 (2000).
- Vogt, K. & Kirschfeld, K. The chemical identity of the chromophores of fly visual pigment. *Naturwissenschaften* **71**, 211–213 (1984).
- Hardie, R. C. The photoreceptor array of the dipteran retina. *Trends Neurosci.* **9**, 419–423 (1986).
- Bloomquist, B. T. et al. Isolation of putative phospholipase C gene of *Drosophila*, *norpA* and its role in phototransduction. *Cell* **54**, 723–733 (1988).
- Scott, K., Becker, A., Sun, Y., Hardy, R. & Zuker, C.  $G_{aq}$  protein function in vivo: genetic dissection of its role in photoreceptor cell physiology. *Neuron* **15**, 919–927 (1995).
- Niemeyer, B. A., Suzuki, E., Scott, K., Jalink, K. & Zuker, C. S. The *Drosophila* light-activated conductance is composed of the two channels TRP and TRPL. *Cell* **85**, 651–659 (1996).
- Putney, J. W. & McKay, R. R. Capacitative calcium entry channels. *BioEssays* **21**, 38–46 (1999).
- Nasi, E., del Pilar Gomez, M. & Payne, R. In *Handbook of Biological Physics* (eds Stavenga, D. G., de Grip, W. J. & Pugh, E. N.) 389–448 (Elsevier, 2000).

20. Acharya, J. K., Jalink, K., Hardy, R. W., Hartenstein, V. & Zuker, C. S. *InsP<sub>3</sub>* receptor is essential for growth and differentiation but not for vision in *Drosophila*. *Neuron* **18**, 881–887 (1997).
21. Raghu, P. *et al.* Normal phototransduction in *Drosophila* photoreceptors lacking an *InsP<sub>3</sub>* receptor gene. *Mol. Cell. Neurosci.* **15**, 429–445 (2000).
22. Smith, D. P. *et al.* Photoreceptor deactivation and retinal degeneration mediated by a photoreceptor-specific protein-kinase-C. *Science* **254**, 1478–1484 (1991).
23. Hardie, R. C. *et al.* Protein kinase C is required for light adaptation in *Drosophila* photoreceptors. *Nature* **363**, 634–637 (1993).
24. Chyb, S., Raghu, P. & Hardie, R. C. Polyunsaturated fatty acids activate the *Drosophila* light-sensitive channels TRP and TRPL. *Nature* **397**, 255–259 (1999).
25. Estacion, M., Sinkins, W. G. & Schilling, W. P. Regulation of *Drosophila* transient receptor potential-like (TrpL) channels by phospholipase C-dependent mechanisms. *J. Physiol.* **530**, 1–19 (2001).
26. Masai, I., Okazaki, A., Hosoya, T. & Hotta, Y. *Drosophila* retinal degeneration A gene encodes an eye-specific diacylglycerol kinase with cysteine-rich zinc-finger motifs and ankyrin repeats. *Proc. Natl Acad. Sci. USA* **90**, 11157–11161 (1993).
27. Raghu, P. *et al.* Constitutive activity of the light-sensitive channels TRP and TRPL in the *Drosophila* diacylglycerol kinase mutant, *rdgA*. *Neuron* **26**, 169–179 (2000).
28. Huang, C. L., Feng, S. Y. & Hilgemann, D. W. Direct activation of inward rectifier potassium channels by PIP<sub>2</sub> and its stabilization by Gβγ. *Nature* **391**, 803–806 (1998).
29. Chuang, H. H. *et al.* Bradykinin and nerve growth factor release the capsaicin receptor from PtdIns(4,5)P<sub>2</sub>-mediated inhibition. *Nature* **411**, 957–962 (2001).
30. Hardie, R. C. *et al.* Calcium influx via TRP channels is required to maintain PIP<sub>2</sub> levels in *Drosophila* photoreceptors. *Neuron* **30**, 149–159 (2001).
31. Li, C. J. *et al.* INAF, a protein required for transient receptor potential Ca<sup>2+</sup> channel function. *Proc. Natl Acad. Sci. USA* **96**, 13474–13479 (1999).
32. Ferreira, P. A. & Pak, W. L. Bovine phospholipase C highly homologous to the *norpA* protein of *Drosophila* is expressed specifically in cones. *J. Biol. Chem.* **269**, 3129–3131 (1994).
33. Womack, K. B. *et al.* Do phosphatidylinositides modulate vertebrate phototransduction? *J. Neurosci.* **20**, 2792–2799 (2000).
34. Scott, K. & Zuker, C. Lights out: deactivation of the phototransduction cascade. *Trends Biochem. Sci.* **22**, 350–354 (1997).
35. Chen, J., Makino, C. L., Peachey, N. S., Baylor, D. A. & Simon, M. I. Mechanisms of rhodopsin inactivation in vivo as revealed by a COOH-terminal truncation mutant. *Science* **267**, 374–377 (1995).
36. Mendez, A. *et al.* Rapid and reproducible deactivation of rhodopsin requires multiple phosphorylation sites. *Neuron* **28**, 153–164 (2000).
37. Vinos, J., Jalink, K., Hardy, R. W., Britt, S. G. & Zuker, C. S. A G protein-coupled receptor phosphatase required for rhodopsin function. *Science* **277**, 687–690 (1997).
38. Byk, T., BarYaacov, M., Doza, Y. N., Minke, B. & Selinger, Z. Regulatory arrestin cycle secures the fidelity and maintenance of the fly photoreceptor cell. *Proc. Natl Acad. Sci. USA* **90**, 1907–1911 (1993).
39. Dolph, P. J. *et al.* Arrestin function in inactivation of G protein-coupled receptor rhodopsin in vivo. *Science* **260**, 1910–1916 (1993).
40. Cook, B. *et al.* Phospholipase C and termination of G-protein-mediated signalling in vivo. *Nature Cell Biol.* **2**, 296–301 (2000).
41. He, W., Cowan, C. W. & Wensel, T. G. RGS9, a GTPase accelerator for phototransduction. *Neuron* **20**, 95–102 (1998).
42. Fesenko, E. E., Kolesnikov, S. S. & Lyubarsky, A. L. Induction by cyclic GMP of cationic conductance in plasma membrane of retinal rod outer segment. *Nature* **313**, 310–313 (1985).
43. Kaupp, U. B. & Koch, K. W. Role of cGMP and Ca<sup>2+</sup> in vertebrate photoreceptor excitation and adaptation. *Annu. Rev. Physiol.* **54**, 153–175 (1992).
44. Hsu, Y. T. & Molday, R. S. Modulation of the cGMP-gated channel of rod photoreceptor cells by calmodulin. *Nature* **361**, 76–79 (1993).
45. Gray, P. & Attwell, D. Kinetics of light-sensitive channels in vertebrate photoreceptors. *Proc. R. Soc. Lond. B* **223**, 379–388 (1985).
46. Kaupp, U. B. Family of cyclic nucleotide gated ion channels. *Curr. Opin. Neurobiol.* **5**, 434–442 (1995).
47. Montell, C. & Rubin, G. M. Molecular characterization of *Drosophila* *trp* locus, a putative integral membrane protein required for phototransduction. *Neuron* **2**, 1313–1323 (1989).
48. Hardie, R. C. & Minke, B. The *trp* gene is essential for a light-activated Ca<sup>2+</sup> channel in *Drosophila* photoreceptors. *Neuron* **8**, 643–651 (1992).
49. Reuss, H., Mojet, M. H., Chyb, S. & Hardie, R. C. In vivo analysis of the *Drosophila* light-sensitive channels, TRP and TRPL. *Neuron* **19**, 1249–1259 (1997).
50. Phillips, A. M., Bull, A. & Kelly, L. E. Identification of a *Drosophila* gene encoding a calmodulin-binding protein with homology to the *trp* phototransduction gene. *Neuron* **8**, 631–642 (1992).
51. Xu, X. Z. S., Chien, F., Butler, A., Salkoff, L. & Montell, C. TRP<sub>γ</sub>, a *Drosophila* TRP-related subunit, forms a regulated cation channel with TRPL. *Neuron* **26**, 647–657 (2000).
52. Harteneck, C., Plant, T. D. & Schultz, G. From worm to man: three subfamilies of TRP channels. *Trends Neurosci.* **23**, 159–166 (2000).
53. Clapham, D. E., Runnels, L. W. & Strubing, C. The *trp* ion channel family. *Nature Rev. Neurosci.* **2**, 387–396 (2001).
54. Hardie, R. C. & Minke, B. Phosphoinositide-mediated phototransduction in *Drosophila* photoreceptors: the role of Ca<sup>2+</sup> and *trp*. *Cell Calcium* **18**, 256–274 (1995).
55. Scott, K., Sun, Y. M., Beckingham, K. & Zuker, C. S. Calmodulin regulation of *Drosophila* light-activated channels and receptor function mediates termination of the light response in vivo. *Cell* **91**, 375–383 (1997).
56. Hardie, R. C. & Mojet, M. H. Magnesium-dependent block of the light-activated and *trp*-dependent conductance in *Drosophila* photoreceptors. *J. Neurophysiol.* **74**, 2590–2599 (1995).
57. Leskov, I. B. *et al.* The gain of rod phototransduction: reconciliation of biochemical and electrophysiological measurements. *Neuron* **27**, 525–537 (2000).
58. Calvert, P. D. *et al.* Membrane protein diffusion sets the speed of rod phototransduction. *Nature* **411**, 90–94 (2001).
59. Neubig, R. Membrane organization in G-protein mechanisms. *FASEB J.* **8**, 939–946 (1994).
60. Sheng, M. & Sala, C. PDZ domains and the organization of supramolecular complexes. *Annu. Rev. Neurosci.* **24**, 1–29 (2001).
61. Tsunoda, S. *et al.* A multivalent PDZ-domain protein assembles signalling complexes in a G-protein-coupled cascade. *Nature* **388**, 243–249 (1997).
62. Montell, C. TRP trapped in fly signalling web. *Curr. Opin. Neurobiol.* **8**, 389–397 (1998).
63. Huber, A. *et al.* The transient receptor potential protein (Trp), a putative store-operated Ca<sup>2+</sup> channel essential for phosphoinositide-mediated photoreception, forms a signaling complex with NorpA, InaC and InaD. *EMBO J.* **15**, 7036–7045 (1996).
64. Li, H. S. & Montell, C. TRP and the PDZ protein, INAD, form the core complex required for retention of the signalplex in *Drosophila* photoreceptor cells. *J. Cell Biol.* **150**, 1411–1422 (2000).
65. Tsunoda, S., Sun, Y., Suzuki, E. & Zuker, C. Independent anchoring and assembly mechanisms of INAD signaling complexes in *Drosophila* photoreceptors. *J. Neurosci.* **21**, 150–158 (2001).
66. Xu, X. Z. S., Choudhury, A., Li, X. L. & Montell, C. Coordination of an array of signaling proteins through homo- and heteromeric interactions between PDZ domains and target proteins. *J. Cell Biol.* **142**, 545–555 (1998).
67. Montell, C. Visual transduction in *Drosophila*. *Annu. Rev. Cell Dev. Biol.* **15**, 231–268 (1999).
68. vanHuizen, R. *et al.* Two distantly positioned PDZ domains mediate multivalent INAD-phospholipase C interactions essential for G protein-coupled signaling. *EMBO J.* **17**, 2285–2297 (1998).
69. Huber, A., Sander, P., Bahner, M. & Paulsen, R. The TRP Ca<sup>2+</sup> channel assembled in a signaling complex by the PDZ domain protein INAD is phosphorylated through the interaction with protein kinase C (cPKC). *FEBS Lett.* **425**, 317–322 (1998).
70. Liu, M. Y., Parker, L. L., Wadzinski, B. E. & Shieh, B. H. Reversible phosphorylation of the signal transduction complex in *Drosophila* photoreceptors. *J. Biol. Chem.* **275**, 12194–12199 (2000).
71. He, W. *et al.* Modules in the photoreceptor RGS9-1/G<sub>βSL</sub> GTPase-accelerating protein complex control effector coupling, GTPase acceleration, protein folding, and stability. *J. Biol. Chem.* **275**, 37093–37100 (2000).
72. Schwarzer, A., Schaaf, H. & Bauer, P. J. Binding of the cGMP-gated channel to the Na/Ca-K exchanger in rod photoreceptors. *J. Biol. Chem.* **275**, 13448–13454 (2000).
73. Korsch, H. G. *et al.* Interaction of glutamic-acid-rich proteins with the cGMP signalling pathway in rod photoreceptors. *Nature* **400**, 761–766 (1999).
74. Lamb, T. D. Gain and kinetics of activation in the G-protein cascade of phototransduction. *Proc. Natl Acad. Sci. USA* **93**, 566–570 (1996).
75. Minke, B. & Hardie, R. C. in *Handbook of Biological Physics* (eds Stavenga, D. G., de Grip, W. J. & Pugh, E. N.) 449–525 (Elsevier, 2000).
76. Pak, W. L., Ostroy, S. E., Deland, M. C. & Wu, C.-F. Photoreceptor mutant of *Drosophila*: is protein involved in intermediate steps of phototransduction? *Science* **194**, 956–959 (1976).
77. Scott, K. & Zuker, C. S. Assembly of the *Drosophila* phototransduction cascade into a signalling complex shapes elementary responses. *Nature* **395**, 805–808 (1998).
78. Bahner, M., Sander, P., Paulsen, R. & Huber, A. The visual G protein of fly photoreceptors interacts with the PDZ domain assembled INAD signaling complex via direct binding of activated G<sub>αq</sub> to phospholipase C<sub>β</sub>. *J. Biol. Chem.* **275**, 2901–2904 (2000).
79. Tsunoda, S. & Zuker, C. S. The organization of INAD-signaling complexes by a multivalent PDZ domain protein in *Drosophila* photoreceptor cells ensures sensitivity and speed of signaling. *Cell Calcium* **26**, 165–171 (1999).
80. Oberwinkler, J. & Stavenga, D. G. Calcium transients in the rhabdomeres of dark- and light-adapted fly photoreceptor cells. *J. Neurosci.* **20**, 1701–1709 (2000).
81. Wong, F. Adapting bump model for ventral photoreceptors of *Limulus*. *J. Gen. Physiol.* **79**, 1089–1114 (1982).
82. Kirschfeld, K. & Vogt, K. Calcium ions and pigment migration in fly photoreceptors. *Naturwissenschaften* **67**, 516–517 (1980).
83. Stavenga, D. G. Insect retinal pigments: spectral characteristics and physiological functions. *Prog. Retinal Eye Res.* **15**, 231–259 (1996).
84. Lamb, T. D. & Pugh Jr A quantitative account of the activation steps involved in phototransduction in amphibian photoreceptors. *J. Physiol.* **449**, 719–758 (1992).
85. Miklos, G. L. G. & Rubin, G. M. The role of the genome project in determining gene function: insights from model organisms. *Cell* **86**, 521–529 (1996).
86. Hardie, R. C. Whole-cell recordings of the light-induced current in *Drosophila* photoreceptors: evidence for feedback by calcium permeating the light sensitive channels. *Proc. R. Soc. Lond. B* **245**, 203–210 (1991).
87. Ranganathan, R., Harris, G. L., Stevens, C. F. & Zuker, C. S. A *Drosophila* mutant defective in extracellular calcium-dependent photoreceptor deactivation and rapid desensitization. *Nature* **354**, 230–232 (1991).
88. Whitlock, G. G. & Lamb, T. D. Variability in the time course of single photon responses from toad rods: termination of rhodopsin's activity. *Neuron* **23**, 337–351 (1999).
89. Shieh, B. H. & Zhu, M. Y. Regulation of the TRP Ca<sup>2+</sup> channel by INAD in *Drosophila* photoreceptors. *Neuron* **16**, 991–998 (1996).
90. Chevesich, J., Kreuz, A. J. & Montell, C. Requirement for the PDZ domain protein, INAD, for localization of the TRP store-operated channel to a signaling complex. *Neuron* **18**, 95–105 (1997).
91. Laughlin, S. B., van Steveninck, R. D. & Anderson, J. C. The metabolic cost of neural information. *Nature Neurosci.* **1**, 36–41 (1998).
92. Warr, C. G. & Kelly, L. E. Identification and characterization of two distinct calmodulin-binding sites in the TrpL ion-channel protein of *Drosophila melanogaster*. *Biochem. J.* **314**, 497–503 (1996).
93. Zhang, Z. *et al.* Activation of Trp3 by inositol 1,4,5-trisphosphate receptors through displacement of inhibitory calmodulin from a common binding domain. *Proc. Natl Acad. Sci. USA* **98**, 3168–3173 (2001).
94. Niemeyer, B. A., Bergs, C., Wissenbach, U., Flockerzi, V. & Trost, C. Competitive regulation of Ca<sup>2+</sup>-like-mediated Ca<sup>2+</sup> entry by protein kinase C and calmodulin. *Proc. Natl Acad. Sci. USA* **98**, 3600–3605 (2001).
95. Fleig, A. *et al.* Calcium-entry second messenger function for ADP-ribose via gating of the LTRPC2 cation channel. *Biophys. J.* **80**, 2749 (2001).
96. Runnels, L. W., Yue, L. X. & Capham, D. E. TRP-PLIK, a bifunctional protein with kinase and ion channel activities. *Science* **291**, 1043–1047 (2001).

## Acknowledgements

We thank R. Paulsen, T. D. Lamb and R. Ranganathan for constructive comments on the manuscript. Apologies to colleagues whose work we failed to cite owing to space limitations.

# Molecular basis of mechanosensory transduction

Peter G. Gillespie & Richard G. Walker

Oregon Hearing Research Center and Vollum Institute, Oregon Health & Science University, Portland, Oregon 97201, USA  
(e-mail: gillespp@ohsu.edu)

**Mechanotransduction — a cell's conversion of a mechanical stimulus into an electrical signal — reveals vital features of an organism's environment. From hair cells and skin mechanoreceptors in vertebrates, to bristle receptors in flies and touch receptors in worms, mechanically sensitive cells are essential in the life of an organism. The scarcity of these cells and the uniqueness of their transduction mechanisms have conspired to slow molecular characterization of the ensembles that carry out mechanotransduction. But recent progress in both invertebrates and vertebrates is beginning to reveal the identities of proteins essential for transduction.**

**M**echanical forces impinge on us from all directions, transmitting valuable information about the external environment. Mechanosensory cells transduce these mechanical forces and transmit this sensory information to the brain. Hearing, touch, sense of acceleration — each informs us about what is nearby and how we are moving relative to our surroundings. An organism detects sensory information with a variety of cells that respond to force. Although different structurally, hair cells within our ear, cutaneous mechanoreceptors of our skin, and invertebrate

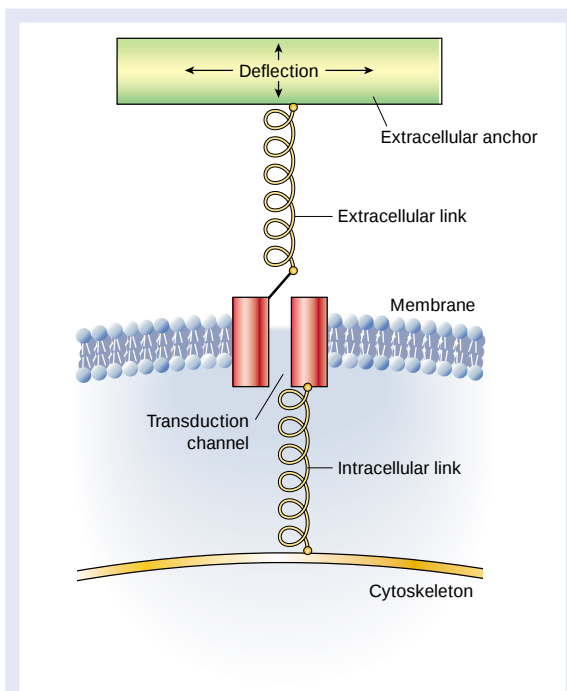
mechanoreceptors share many mechanistic features; whether mutual molecular mechanisms underlie these similar transduction mechanisms remains to be determined. Here we review mechanisms of mechanosensory transduction by invertebrates and vertebrates, highlighting the increasing molecular understanding of transduction in each system.

## General features of mechanotransduction

As with most sensory systems, mechanosensory cells place a premium on speed and sensitivity. A common theme is for mechanical forces to be directed to specific ion channels, which can open rapidly and amplify the signal by permitting entry of large numbers of ions. Mechanical forces can also affect intracellular events in cells — such as gene transcription — directly through the cell surface and cytoskeleton, although such mechanisms typically are not used for rapid sensory transduction.

Speed requires that mechanical forces be funnelled directly to transduction channels, without intervening second messengers. Sensitivity requires that the maximal amount of stimulus energy be directed to the transduction channel. A general model — borrowed from worm touch receptors<sup>1,2</sup> and hair cells<sup>3</sup> — that applies to many mechanosensory transduction systems is illustrated in Fig. 1; its key feature is a transduction channel that detects deflection of an external structure relative to an internal structure, such as the cytoskeleton. Such a deflection could take the form of deformation of the skin, oscillation of a hair cell's hair bundle, or vibration of a fly's bristle. Deflection changes tension in all elements of the system, and the transduction channel responds by changing its open probability.

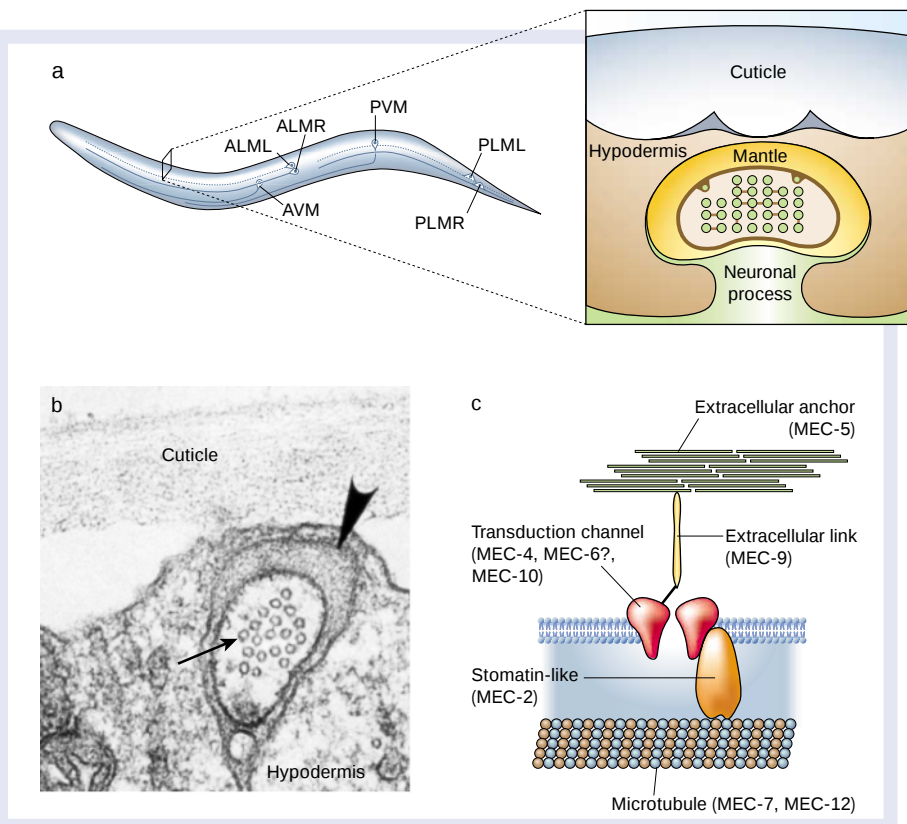
This general mechanism also explains other features that are common to mechanosensory cells, such as adaptation. Vestibular hair cells, for example, must reject the constant 1-g gravitational force to detect stimuli one-millionth that size<sup>4</sup>. During adaptation to a sustained stimulus, where there is a constant relationship between external and internal structures, tension applied to the transduction channel declines through a readjustment of the machinery. Adaptation could occur through deformation of the external coupling structure, lengthening of the external or internal anchors, slipping of the internal anchor relative to the cytoskeleton, or changes within the channel



**Figure 1** General features of mechanosensory transduction. A transduction channel is anchored by intracellular and extracellular anchors to the cytoskeleton and to an extracellular structure to which forces are applied. The transduction channel responds to tension in the system, which is increased by net displacements between intracellular and extracellular structures.



**Figure 2** *C. elegans* touch-receptor structure and transduction model. **a**, View of *C. elegans* showing positions of mechanoreceptors. AVM, anterior ventral microtubule cell; ALML/R, anterior lateral microtubule cell left/right; PVM, posterior ventral microtubule cell; PLML/R, posterior lateral microtubule cell left/right. **b**, Electron micrograph of a touch-receptor neuron process. Mechanotransduction may ensue with a net deflection of the microtubule array relative to the mantle, a deflection detected by the transduction channel. Arrow, 15-protofilament microtubules; arrowhead, mantle. Modified from ref. 3. **c**, Proposed molecular model for touch receptor. Hypothetical locations of *mec* proteins are indicated.



itself. Appropriately, all of these structures and interactions are amenable to molecular characterization.

### Invertebrate mechanoreceptor models

Interesting in their own right, invertebrate mechanoreceptors have also attracted considerable attention because they can be readily approached with genetics. The two most utilized invertebrate models — the nematode *Caenorhabditis elegans* and fruitfly *Drosophila melanogaster* — have yielded candidate transduction channels, as well as possible accessory molecules.

Other kingdoms possess mechanoreceptors too. Although not a mechanosensory transduction channel in the sense we discuss here, the cloned and reconstituted MscL channel of *Escherichia coli* responds unequivocally to membrane tension in the absence of other proteins<sup>5</sup>. The discovery of these channels in archaeobacteria<sup>6</sup> highlights the ancient nature of mechanotransduction and its critical role in all cells. Although no MscL homologues have been identified in eukaryotes, the ability to determine the functional consequences of structural perturbations should permit an unparalleled view of a mechanically gated channel.

### Touch mechanotransduction in *C. elegans*

A simple and ingenious genetic screen identified *C. elegans* mutants (*mec* mutants) that were defective in mechanosensation<sup>7,8</sup>. Mutant worms that responded inappropriately or not at all to the simple touch of an eyelash were selected and most of the responsible genes have been identified. Although some of the gene products participate in development of the six mechanosensory touch neurons in the worm<sup>9,10</sup>, most of the gene products seem to constitute a transduction machinery (Fig. 2). The identities and interactions of cloned mutant genes suggest a mechanoreceptive complex, with a mechanically sensitive ion channel connected to both intracellular cytoskeletal components and extracellular matrix proteins (including the mantle surmounting the touch-receptor process) to form a transduction apparatus (reviewed in refs 1, 2).

Consistent with the general model for a mechanotransduction apparatus (Fig. 1), which requires an intra- and extracellularly

tethered channel, several of the *mec* genes encode components of the cytoskeleton, the extracellular matrix, or the links to them (Fig. 2). The genes *mec-7* and *mec-12* encode  $\beta$ - and  $\alpha$ -tubulins<sup>11,12</sup>, which constitute the 15-protofilament microtubules in sensory dendrites of touch neurons. Perhaps serving as a linker between these microtubules and a transduction channel, the MEC-2 protein is expressed exclusively by touch-receptor neurons and shows homology to stomatin, an integral membrane protein found in many cell types, which associates with the cytoskeleton and is involved in ionic homeostasis of the cell<sup>13</sup>. MEC-5 is a specialized collagen-like molecule secreted by the hypodermal cells that surround the receptor process and may be an extracellular anchor for the transduction apparatus<sup>14</sup>. The multidomain MEC-9 protein is expressed and secreted by the sensory neurons<sup>14</sup> and interacts genetically with MEC-5 and the proposed transduction channel MEC-4 (ref. 15), suggesting that they function together as part of the extracellular linkages of the transduction machinery.

### Transduction by DEG/ENaCs

The most intriguing set of genes to come from the screens for *mec* mutants were the degenerins (DEGs), which encode ion channels related to vertebrate epithelial sodium channels (ENaCs) responsible for  $\text{Na}^+$  adsorption (reviewed in ref. 16; see Box 1). Most mutations in the *C. elegans* DEG genes *mec-4* and *mec-10* render the animal insensitive to light touch, whereas dominant gain-of-function mutations in *mec-4* and *mec-10*, as well as *unc-8* and *deg-1*, cause degeneration of the touch neurons (hence the familial name), probably by causing the channels to be open constantly<sup>17</sup>. Because of their critical role in the touch-receptor neurons, DEGs have been proposed to act as mechanosensory transduction channels or subunits thereof<sup>18</sup>. Attempts to elicit mechanically induced currents from heterologous cells expressing these channels have been unsuccessful<sup>19</sup>, although suction-induced currents have been observed in patches containing the related vertebrate ENaC  $\alpha$ -subunit<sup>20</sup>. It is possible that DEGs might require specific interactions with both intracellular and extracellular tethers and even other channel subunits for mechanical gating. Confirmation of these molecules as transduction

# Box 1

## DEG/ENaC channels

Members of the growing DEG/ENaC superfamily traverse the membrane twice and have intracellular N and C termini; a 33-residue portion of the N terminus is conserved among family members (reviewed in ref. 16). This domain, which shows similarity to thiol proteases, including the conservation of a catalytic histidine residue, is found near the first transmembrane region and may be important in subunit interactions. A large extracellular loop contains two or three cysteine-rich regions and is followed by the second transmembrane segment, which is thought to function as the channel's pore.

**Degenerins.** These include *mec-4* and *mec-10*, present in *C. elegans* touch receptors, as well as *unc-8*, *unc-105* and *deg-1*. Interacting genetically with type IV collagen in *C. elegans* muscle, *unc-105* might mediate stretch sensitivity in muscle<sup>62</sup>. Certain mutations in these genes cause dominant degenerative cell death, perhaps by constitutive ion-channel activation.

**Epithelial sodium channels (ENaCs).** Also called amiloride-sensitive sodium channels for their sensitivity to this drug, the ENaCs are responsible for Na<sup>+</sup> resorption in many epithelia, such as the kidney, distal colon, secretory glands and respiratory airways. In addition, the sense of salty taste is mediated by ENaC channels in the fungiform papillae of the tongue (see review in this issue by Lindemann, pages 219–225). The C terminus of ENaCs contains a

proline-rich PY domain that mediates regulation of the channels by Nedd4, a ubiquitin-protein ligase that binds specifically to the PY domain. These molecules form heteromultimeric channels with some combination of  $\alpha$ -,  $\beta$ -,  $\gamma$ - and  $\delta$ -subunits likely in either a tetrameric or octomeric configuration.

**Acid-sensing ion channels (ASICs).** Found in a variety of tissues, ASICs are activated by extracellular protons and may mediate pain induced by acidosis. ASIC1 is activated at relatively high pH levels and inactivates rapidly; there are two variants, ASIC1a and ASIC1b. ASIC2 also has two variants — ASIC2a, known additionally as BNC1 (brain sodium channel 1, sometimes abbreviated BNaC1), and ASIC2b, also known as BNC2 or ASIC- $\beta$ . The BNCs have also been labelled MDEGs (mammalian degenerins), a phylogenetically inappropriate term. ASIC3 was originally named DRASIC (dorsal-root acid-sensing ion channel), although this channel is actually found in many tissues.

**Drosophila family members.** Of the ~14 members of DEG/ENaC gene family in the fly genome, only two have been characterized: *ripped pocket* (*rpk*) and *pickpocket* (*ppk*). Because PPK is expressed in peripheral neurons thought to mediate mechanosensation, this channel is a strong transduction-channel candidate.

channels, however, awaits electrophysiological recording of receptor currents from mutant touch neurons.

Given the importance of DEG/ENaC family members in *C. elegans* mechanoreceptors, a role for these channels in the mechanical senses of other organisms seems likely. For example, there are at least 14 members of DEG/ENaC gene family in the fly genome, two of which, *ripped pocket* (*rpk*) and *pickpocket* (*ppk*), have been characterized<sup>21</sup>. There are currently no described mutants in either *rpk* or *ppk* with which to evaluate the role of these channels in the fly, although restricted expression of PPK to mechanosensory neurons is suggestive of a role in mechanosensation.

Vertebrates also have a large complement of DEG/ENaC channels (Box 1); the human genome encodes at least a dozen such genes. More recent research on these channels has implicated them in various mechanosensory modalities. For example, the  $\gamma$ -subunit of ENaC is present in sensory endings of baroreceptors, the mechanosensory neurons that sense blood pressure<sup>22</sup>, and mechanically stimulated baroreceptor activity is blocked by benzamil, an amiloride analogue that also blocks ENaCs<sup>22</sup>. Another member of the DEG/ENaC superfamily, brain sodium channel 1 $\alpha$  (BNC1 $\alpha$ ), a splice variant of BNC1, is expressed in dorsal root ganglia and transported to a number of cutaneous mechanosensory terminals<sup>23</sup>. BNC1 $\alpha$  is found in most identified cutaneous mechanoreceptors, and not in endings specialized for other sensory modalities, indicating a particular role in touch.

To evaluate the role of BNC1 in touch sensation, mechanoreceptor responses were examined in mice whose BNC1 gene had been deleted<sup>24</sup>. In homozygous mutant mice, rapidly adapting neurons fired only about half the number of action potentials as wild-type mice in response to a 20- $\mu$ m stimulus. Furthermore, slowly adapting neurons from BNC1-null mutants showed a modest decrease in sensitivity. One might expect that deleting a transduction channel or one of its subunits from the transduction complex might produce a more profound phenotype — BNC1-null neurons still responded to stimuli and there were no obvious behavioural deficits — but other members of this family may have compensated for the loss of BNC1. Indeed, the  $\beta$ - and  $\gamma$ -subunits of ENaC have been recently localized

to mechanoreceptor terminals in the skin<sup>25,26</sup>. Unfortunately, mice bearing targeted disruptions of these two genes die of defects in electrolyte metabolism within a few days of birth (reviewed in ref. 27), making analysis of their mechanoreceptor function difficult.

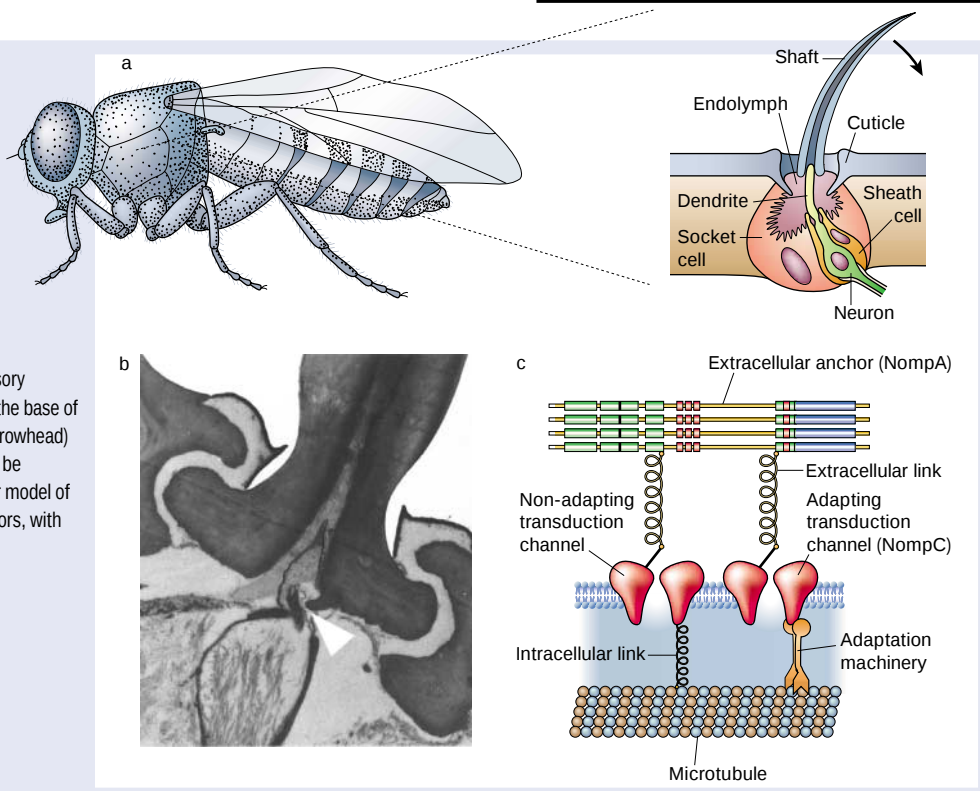
## Drosophila mechanotransduction

When Kernan and colleagues screened for mechanosensory mutants in fruitflies using an adaptation of the *C. elegans* screen<sup>28</sup>, they put *Drosophila* on the mechanosensory map. Renowned and increasingly powerful genetic tools coupled with the ability to record mechanosensory receptor potentials and currents from bristles make *Drosophila* a consummate model system for dissecting mechanosensation. The fly mechanosensory system comprises two sets of mechanoreceptors. Type I sensory organs have one to three sensory neurons, each with a single ciliated sensory dendrite, supported by accessory cells (Fig. 3), whereas type II mechanoreceptor neurons have multiple nonciliated dendrites and no accessory cells. Type I mechanoreceptors include bristle mechanoreceptors, chordotonal organs such as the Johnston's organ (the fly's antennal hearing apparatus) and campaniform sensilla, which are featured prominently in halteres, the club-shaped mechanoreceptive structures that relay information about wing beat. Because the large bristles are hollow and contain a conductive high-K<sup>+</sup> endolymph (similar to the endolymph of vertebrate hearing and vestibular organs) that also bathes the apical surface of the sensory epithelium, electrical access to the tiny sensory dendrite can be gained by snipping the bristle and placing a recording (and stimulation) electrode over its cut end<sup>28,29</sup>. It is possible to clamp the potential across this sensory epithelium and record mechanically gated transduction currents<sup>29</sup>.

To test whether the larval mutants identified by the behavioural screen were defective in the transduction pathway, mechanoreceptor potentials were recorded from mutant adult sensory bristles and compared with those from the bristles of wild-type flies. A number of the larval touch-insensitive mutants developed into profoundly uncoordinated adults and showed either a reduced mechanoreceptor potential (*remp* mutants) or no mechanoreceptor potential (*nomp* mutants) at all. Evolution has apparently conserved the transduction

**Figure 3** *Drosophila* bristle-receptor model.

**a**, Lateral view of *D. melanogaster* showing the hundreds of bristles that cover the fly's cuticle. The expanded view of a single bristle indicates the locations of the stereotypical set of cells and structures associated with each mechanosensory organ. Movement of the bristle towards the cuticle of the fly (arrow) displaces the dendrite and elicits an excitatory response in the mechanosensory neuron. **b**, Transmission electron micrograph of an insect mechanosensory bristle showing the insertion of the dendrite at the base of the bristle. The bristle contacts the dendrite (arrowhead) so that movement of the shaft of the bristle will be detected by the neuron. **c**, Proposed molecular model of transduction for ciliated insect mechanoreceptors, with the locations of NompC and NompA indicated.



mechanism of various mechanosensory modalities in flies, as almost all of the uncoordinated mutants identified in this screen are also deaf<sup>30</sup>.

So far, only two of the mutant genes have been identified molecularly, *nompA*<sup>31</sup> and *nompC*<sup>29</sup>. Like the results from the *C. elegans* screen, the screen from flies has produced an extracellular-matrix protein and an ion channel (Fig. 3). The mechanosensory gene *nompA* encodes a modular protein, secreted by the supporting cells and localized to the extracellular dendritic cap-like structures of type I mechanoreceptors and some chemoreceptors. NompA probably serves as an extracellular mechanical link between the sensory dendrites of type I mechanoreceptors and their associated cuticular structures. It shows similarity to a number of proteins, including the tectorins<sup>32</sup>, extracellular molecules that transmit mechanical stimuli to mechanoreceptive hair cells in the vertebrate auditory system.

The potency of coupling electrophysiology with traditional *Drosophila* genetics was demonstrated in the cloning of the mechanosensory gene *nompC*<sup>29</sup>. Although three alleles of *nompC* were severely uncoordinated and showed near-complete abrogation of transduction current, a fourth allele had near wild-type amplitude of response, but noticeably faster adaptation than controls. This indicated that the NompC protein would be intimately involved with the transduction and adaptation process.

The *Drosophila* *nompC* gene and a closely related *C. elegans* homologue encode a new and unusual member of the transient receptor potential (TRP) family of ion channels (see review in this issue by Hardie and Raghu, pages 186–193). The predicted protein of 1,619 amino acids can be divided into two regions: first, an unusual ~1,150-amino-acid amino terminus composed of 29 ankyrin repeats, and second, a carboxy-terminal segment with six transmembrane domains, a predicted pore loop, and sequence similarity to TRP channels. Ankyrin repeats are 33-residue motifs with a conserved backbone and variable residues that mediate specific protein–protein interactions<sup>33</sup>. Ankyrin repeats are found in numerous proteins of widely varying functions and subcellular locations<sup>34</sup>. Other members of the TRP channel family have between one and four ankyrin repeats of unknown function; that NompC bears 29 such domains suggests that it interacts either

very strongly with a small number of partners or less avidly with many molecules.

Several lines of evidence led to the conclusion that NompC serves as a mechanosensory transduction channel. (1) NompC shows similarity with other sensory transduction channels in its primary sequence. (2) Loss-of-function *nompC* mutants show loss of almost all transduction current, whereas a point mutation in NompC changes the adaptation profile of the transduction current. (3) The *Drosophila* *nompC* gene is expressed specifically in mechanosensory organs. (4) A *C. elegans* NompC–GFP fusion construct is expressed in putative mechanosensory neurons at the site of transduction. (5) The N terminus of NompC is replete with ankyrin domains, a property one might expect of a mechanically tethered transduction channel.

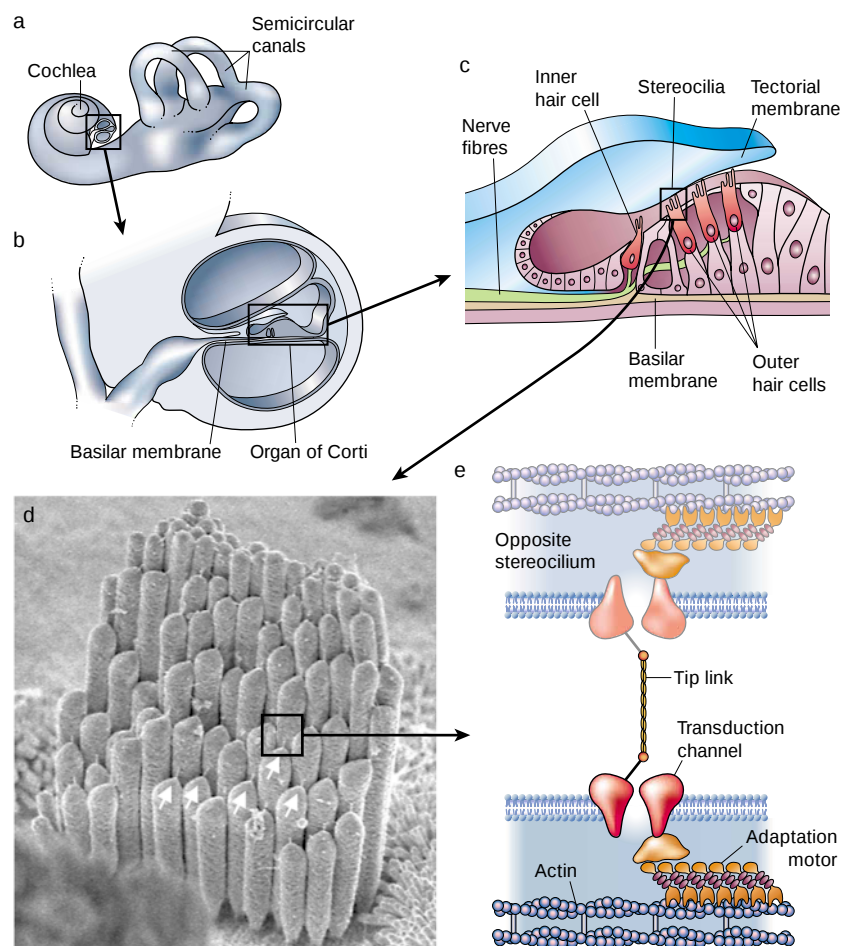
Because opening of the transduction channels in fly mechanoreceptors takes place within 200  $\mu$ s of a mechanical stimulus<sup>29</sup>, fly mechanotransduction seems to take place by direct opening of a mechanically gated conductance without intervening second messengers. Although NompC accounts for ~90% of the transduction current, it is unclear whether the NompC channel is itself mechanically sensitive or instead takes its gating cue from the small mechanically activated current that can be observed in *nompC*-null traces. NompC would then be serving as an adapting amplifier of the true mechanically gated channel.

TRP channels and the closely related cyclic nucleotide-gated and vanilloid-receptor channels are involved or expressed in many sensory systems, including phototransduction, olfaction and taste, as well as in the perception of heat, acid and pain (see accompanying reviews in this issue). So the discovery that a member of this superfamily acts as or influences a transduction channel in mechanosensation should not be unexpected. Indeed, TRP family members have also been implicated in mechanotransduction's sister sense, osmosensation. A *C. elegans* TRP channel, OSM-9, is expressed in dendrites of some ciliated sensory neurons and is required for osmosensation and nose touch sensation<sup>35</sup>. In addition, a new mammalian member of the vanilloid-receptor family is expressed in osmotically sensitive cells and forms osmotically gated channels when expressed in cells<sup>36</sup>.

Cilium-bearing mechanoreceptors in both *C. elegans* and *Drosophila* express NompC, whereas the non-ciliated touch-receptor



**Figure 4** Inner-ear structure and hair-cell transduction model. **a**, Gross view of part of the inner ear. Sound is transmitted through the external ear to the tympanic membrane; the stimulus is transmitted through the middle ear to the fluid-filled inner ear. Sound is transduced by the coiled cochlea. **b**, Cross-section through the cochlear duct. Hair cells are located in the organ of Corti, resting on the basilar membrane. **c**, Sound causes vibrations of the basilar membrane of the organ of Corti; because flexible hair-cell stereocilia are coupled to the overlying tectorial membrane, oscillations of the basilar membrane cause back-and-forth deflection of the hair bundles. **d**, Scanning electron micrograph of hair bundle (from chicken cochlea). Note tip links (arrows). **e**, Proposed molecular model for hair-cell transduction apparatus.



neurons in worms and the type II multiple dendritic neurons of flies instead express members of the DEG/ENaC family. This suggests not only a morphological schism, but also a mechanistic difference between these two kinds of mechanoreceptors. All vertebrate hair cells are endowed with a true cilium, or kinocilium, during the construction of the actin-based hair bundle. Given this odd remnant of ontogeny, it has been speculated that ciliated mechanoreceptors of invertebrates might be related to their once-ciliated vertebrate counterparts. Indeed this relationship extends further to the conservation of molecules that control the development of vertebrate and invertebrate mechanoreceptor organs. Lateral inhibition through signalling of the Notch receptor and its ligand Delta controls the cell-fate patterning of both *Drosophila* mechanoreceptor neurons<sup>37</sup> and vertebrate hair cells<sup>38</sup>. In addition, fly Atonal and its mammalian homologue Math1 also control the specification of cell type. *Drosophila atonal* mutants are devoid of chordotonal organs; similarly, *Math1*-knockout mice generate no hair cells during development of the sensory epithelium<sup>39</sup>. Their signalling is so similar that Math1 can substitute for Atonal in development of *Drosophila* mechanoreceptor organs<sup>40</sup>. Although both signalling modalities are found in other cell types, their control over vertebrate and invertebrate mechanoreceptors, as well as the sharing of specialized sensory structures, suggests that mechanoreceptors evolved prior to the common ancestor of invertebrates and vertebrates and that those mechanoreceptors might share more than just their developmental algorithms.

### Auditory and vestibular transduction by hair cells

Hair cells, the mechanoreceptors of the inner ear, transduce auditory and vestibular stimuli to allow us to hear and sense movements of our

head. Jutting apically from a hair cell, the mechanically sensitive hair bundle bends back and forth in response to stimuli that are directed to it. Auditory stimuli induce a vibration of the structure on which the hair cells sit (the basilar membrane; Fig. 4). Vestibular stimuli cause displacement of acellular structures overlying the hair cells (otolithic membrane in the saccule and utricle, responsible for linear-acceleration detection; cupula in the semicircular canals, responsible for rotational detection), resulting in bundle deflection. An excitatory deflection of a hair bundle (Fig. 4) directly opens transduction channels, which admit cations and depolarize the hair cell. Inhibitory deflections close transduction channels and hyperpolarize the cell. These changes in membrane potential in turn increase (depolarization) or decrease (hyperpolarization) neurotransmitter release from graded synapses on basolateral surfaces of hair cells.

### Transduction mechanism

The gating-spring model<sup>41</sup> successfully describes key biophysical features of mechanical transduction in hair cells. Transduction channels open so fast—within microseconds of a stimulus—that second-messenger cascades cannot have a central role<sup>41</sup>. Instead, elastic gating springs transmit external forces to transduction channels, and the channels' open probability depends on the tension in these springs. Thus, when tension is high, channels spend most of their time open; when tension is low, channels close. Although the original formulation of the model indicated that the open probability of transduction channels in the absence of any gating-spring tension was <0.0001 (ref. 3), recent data from auditory hair cells suggest that the zero-tension open probability is 100-times larger<sup>42</sup>. This debate is relevant to channel identification; higher open probabilities for the channel in the

## Box 2

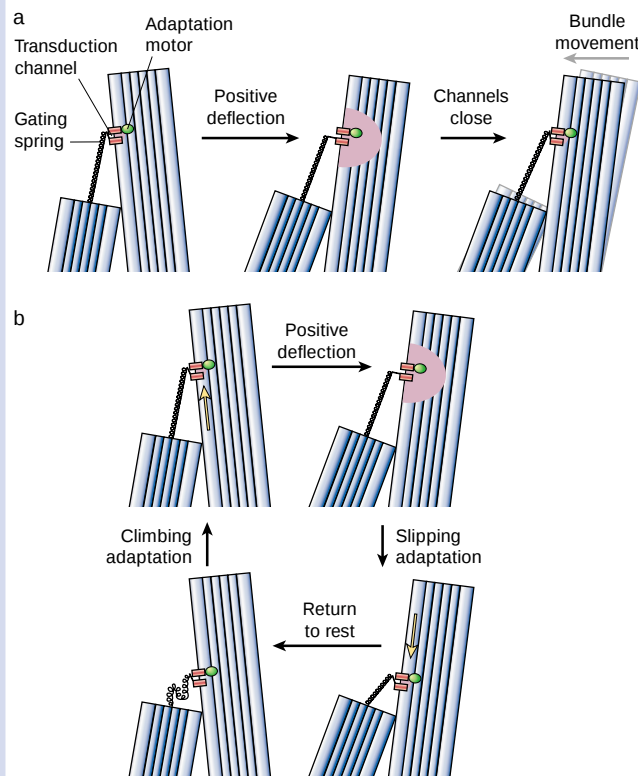
### Adaptation by hair cells

Two forms of adaptation are present in hair cells, each of which can modify the mechanical properties of the bundle (see figure opposite). Fast adaptation is remarkably quick (a millisecond or less in turtle auditory hair cells<sup>63</sup>), requires  $\text{Ca}^{2+}$  (which enters through open transduction channels), and is tuned within individual hair cells, so that cells responding to high sound frequencies adapt more speedily than those tuned to lower frequencies<sup>63</sup>.

Fast adaptation correlates with active hair-bundle movements<sup>64,65</sup>; hair bundles respond to a simple stimulus with a complex, active mechanical response. In hair cells of lower vertebrates, a step stimulus near threshold can initiate a bundle response in the opposite direction with sufficient force to move the bundle back to its original position.

In vestibular hair cells, a distinct mechanism — negative hair-bundle stiffness — permits oscillatory behaviour that is controlled by the mechanism responsible for slow adaptation, the adaptation motor<sup>66</sup>. A bundle with negative hair-bundle stiffness actually moves farther in response to a stimulus than the size of the stimulus itself. The adaptation motor continuously moves the hair bundle into the range of negative stiffness, triggering bundle oscillations that enable much more sensitive detection of small signals.  $\text{Ca}^{2+}$ -dependent channel closure could act synergistically with negative hair-bundle stiffness to power even larger active bundle movements.

Slow adaptation is thought to be mediated by an adaptation motor. Because the force producer of the adaptation motor may be a cluster of myosin molecules<sup>67</sup>, the motor has attracted much attention. The strongest candidate is myosin 1c (Myo1c; formerly known as myosin Ib), which is located near stereociliary tips at tip-link ends<sup>67</sup>. Myosin VIIA (Myo7a) may also be important in adaptation. In outer hair cells that are genetically deficient in this myosin, large displacements are required to begin to open channels, indicating that these cells lack the force generator that maintains resting tension<sup>68</sup>.



**Box 2 Figure** Hair-cell transduction and adaptation. **a**, Transduction and fast adaptation. At rest (left panel), transduction channels spend ~5% of the time open, allowing a modest  $\text{Ca}^{2+}$  entry (pink shading). A positive deflection (middle) stretches the gating spring (drawn here as the tip link); the increased tension propagates to the gate of the transduction channel, and channels open fully. The resulting  $\text{Ca}^{2+}$  flowing in through the channels shifts the channels' open probability to favour channel closure (right). As the gates close, they increase force in the gating spring, which moves the bundle back in the direction of the original stimulus. **b**, Transduction and slow adaptation. Slow adaptation ensues when the motor (green oval) slides down the stereocilium (lower right), allowing channels to close. After the bundle is returned to rest (lower left), gating-spring tension is very low; adaptation re-establishes tension and returns the channel to the resting state.

absence of external force suggest that channel conductance might be detectable when expressed in a heterologous cell type.

The anatomical correlate of the gating spring has long been assumed to be the tip link, a fine extracellular filament that connects adjacent stereocilia along the sensitivity axis<sup>43</sup>. Tip links are in position to be stretched by excitatory stimuli and slackened by inhibitory ones (Fig. 4). But high-resolution structural characterization of the tip link has indicated that it might be too stiff to account for the gating spring. Instead, the elastic element within the bundle might reside elsewhere, such as the intracellular filaments at the basal insertion of the tip link<sup>44</sup>.

#### Transduction-channel properties

One approach to determining the identity of molecules required for transduction has been to extensively characterize transduction and adaptation, then suggest candidate molecules for the channel or motor. But the transduction channel, a nonselective cation channel<sup>45</sup>, has few properties that distinguish it from other channels. Its conductance is large, ~100 pS (ref. 46), and its permeability to  $\text{Ca}^{2+}$  is substantial<sup>47</sup>. The channel is blocked by relatively low concentra-

tions of aminoglycoside antibiotics<sup>48</sup>, amiloride and its derivatives<sup>49</sup>,  $\text{La}^{3+}$  (ref. 50), tubocurarine<sup>51</sup>, and  $\text{Ca}^{2+}$ -channel antagonists such as D-600 (ref. 52) and nifedipine<sup>53</sup>. Unfortunately, this inhibition spectrum does not obviously match any other known channel type and the affinities of these blockers do not allow their use in biochemical isolation of transduction channels from inner-ear tissue.

Early hopes that the transduction channel might incorporate the  $\alpha$ -subunit of the ENaC family were dashed by the persistence of normal transduction in mice lacking this subunit<sup>54</sup>. But the DEG/ENaC family is large, so the hair-cell transduction channel may yet be identified from this family. An alternative candidate is the  $\text{P2X}_2$  receptor, a purinergic-receptor channel that is located in hair bundles<sup>55</sup>, although expression of  $\text{P2X}_2$  is low<sup>56</sup> at a time when hair cells transduce vigorously<sup>57</sup>. Furthermore, the transduction channel and hair-cell  $\text{P2X}_2$  receptor show distinct patterns of blockade by inhibitors<sup>51</sup>. Another possible source for the transduction channel is the TRP family, which includes at least two channels definitely involved in mechanosensation (NompC in flies and OSM-9 in worms); the conductance properties of this family certainly could accommodate those of the hair-cell transduction channel<sup>58</sup>.

### Box 3

#### Essential molecules for hair cells

Hair-bundle proteins identified by genetics or other means can be assembled into an incomplete molecular model (see figure). Hair bundles contain more than 30 major proteins<sup>69</sup>, however, indicating many more molecules remain to be identified.

**Usher genes.** At least ten different genes can cause Usher syndrome, with varying clinical impact. Usher 1, caused by mutations at six or more loci, is the most clinically severe form of the disease. The first of these genes identified, that responsible for Usher 1B, was *MYO7A*<sup>70</sup>. Hair bundles are disarrayed when *Myo7a* from mouse<sup>71</sup> or zebrafish<sup>72</sup> is mutated, which supports a role for *Myo7a* in anchoring ankle links, a class of stereociliary crosslinks near the base of the stereocilia<sup>73</sup>. A new member of the cadherin family of  $\text{Ca}^{2+}$ -dependent cell-adhesion molecules, *Cdh23*, is responsible for Usher 1D<sup>74</sup>; because stereocilia adhesion is disrupted in the corresponding mouse mutant, *waltzer*, *Cdh23* probably forms one of the classes of stereociliary crosslinks<sup>75</sup>. The gene mutated in Usher 1F<sup>76,77</sup> and the mouse deafness model *Ames waltzer*<sup>78</sup> encodes another member of the cadherin family, protocadherin-15 (*Pcd15*). *Ames waltzer* mice have disorganized stereocilia, suggesting that *Pcd15* might crosslink adjacent stereocilia. Mutations in the genes encoding *Myo7a*, *Pcd15* and *Cdh23* produce an identical clinical phenotype, indicating that these proteins may be part of the ankle-link complex. A new protein — harmonin or *USH1C* — is mutated in Usher 1C; its multiple PDZ domains indicate a role in assembling a complex of proteins, although it might also assemble with ankle links. Usher 2 has three identified loci and is characterized by a modest hearing loss, normal balance and late-onset retinitis pigmentosa. *USH2A*, a molecule with multiple domains shared by extracellular matrix/adhesion molecules, is mutated in Usher 2A<sup>79</sup>.

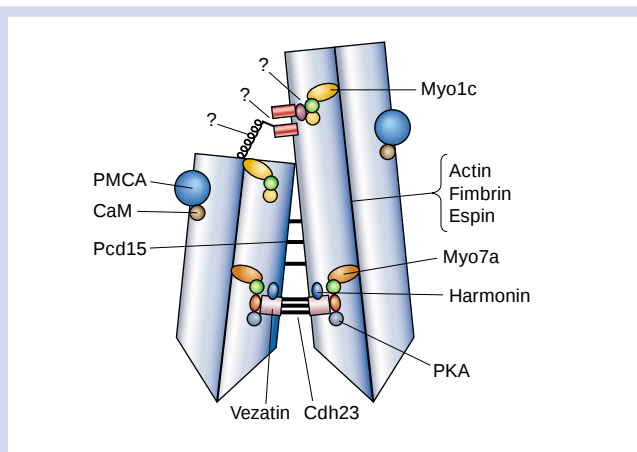
**Myo7a-interacting proteins.** Two proteins have been identified by two-hybrid screening with *Myo7a* bait. One is the RII subunit of the cyclic AMP-dependent protein kinase<sup>80</sup>, which coincides with an earlier observation that the resting open probability of the transduction channel declines in response to agents that raise cAMP levels<sup>63</sup>. The other is a new protein named vezatin<sup>81</sup>, which localizes in hair cells to the ankle-link region and, in tissue-culture cells, interacts with cadherin-based cellular junctions. This protein most likely forms the transmembrane link between *Myo7a* and the ankle links.

**Other myosin isoforms.** *Myo6* and *Myo15* may participate in hair-bundle formation and maintenance. *Myo6*, a negative-end-directed motor, localizes to the cuticular plate and basal insertions of the stereocilia; it may provide outward force on stereocilia or anchor the stereociliary membrane<sup>73,82</sup>. Mice lacking *Myo15* have abnormally short stereocilia, indicating that this isoform might participate in actin-filament elongation (as do yeast myosin I isoforms<sup>83</sup>).

**$\text{Ca}^{2+}$  control in stereocilia.** Because  $\text{Ca}^{2+}$  entry through open transduction channels is essential for fast and slow adaptation and may regulate tip-link assembly<sup>84</sup>, proteins that control  $\text{Ca}^{2+}$  levels in stereocilia should be essential for hair-cell function. A conventional knockout approach<sup>85</sup>, as well as identification of the mouse deafness mutants *deafwaddler* (*dfw*)<sup>86</sup> and *Wriggle mouse sagami* (*wri*)<sup>87</sup>,

showed that mice require isoform 2 of the plasma-membrane  $\text{Ca}^{2+}$ -ATPase (PMCA) for proper auditory and vestibular function. In all hair cells, the only isoform found in the bundle is PMCA2a<sup>88</sup>, which can be present at the remarkably high density of 2,000 molecules per  $\mu\text{m}^2$  (ref. 89). Calmodulin is present in hair bundles at high concentrations<sup>90,91</sup>, where it regulates PMCA, myosins and presumably other bundle molecules<sup>92</sup>. A gene that interacts genetically with *dfw*, *modifier of deafwaddler* (*mdfw*)<sup>93</sup>, seems to be a *Cdh23* allele<sup>75</sup>. Mice heterozygous for *dfw* become deaf only when homozygous for *mdfw*; a possible explanation is that the level of intracellular  $\text{Ca}^{2+}$  in the stereocilia — presumably higher when only one PMCA2a allele is present — affects the function of this candidate *Cdh23* allele. Furthermore, a gene that imparts age-related hearing loss, *ahl1*, also seems to be a *Cdh23* allele.

**Structural proteins.** A properly formed hair bundle is essential for mechanotransduction. Actin has long been known to be the main cytoskeletal element of hair bundles<sup>94</sup>, and fimbrin was identified as an important crosslinker over a decade ago<sup>95</sup>. More recently, espin was identified as a second crosslinker of stereocilia, responsible for the *jerker* mouse deafness mutation<sup>96</sup>. In addition, a human homologue of the fly protein diaphanous is mutated in the nonsyndromic deafness *DFNA1* (ref. 97). Diaphanous is a ligand for the actin-binding protein profilin and is a target for regulation by Rho, which regulates cytoskeletal assembly in many cell types.



**Box 3 Figure** Schematic illustration of identified hair-bundle proteins showing hypothetical locations of molecules implicated in stereocilia function. *Myo7a*, vezatin, *Cdh23* and PKA may form the ankle-link complex. *Pcd15* presumably also interconnects stereocilia. *Myo1c* may carry out adaptation, whereas PMCA maintains a low  $\text{Ca}^{2+}$  concentration. Actin, fimbrin and espin have structural roles; not shown is *DFNA1*, which may help form the cytoskeleton. Calmodulin regulates several enzymes within the bundle, including PMCA, *Myo1c* and *Myo7a*.

#### Adaptation by hair cells

Like all other sensory receptors, hair cells respond to sustained stimuli by adapting, which restores their sensitivity to threshold deflections. The mechanisms responsible for adaptation also set the resting tension in gating springs; the resulting open probability enables maximal sensitivity and high-frequency signal transmission<sup>59</sup>. Properties of adaptation vary substantially from preparation to preparation, and recent data suggest that this arises from differences in hair cells in different organs. Two distinct  $\text{Ca}^{2+}$ -dependent forms of adaptation operate simultaneously in hair cells that have

been examined closely; one form is fast and involves the transduction channel directly, whereas the other is slower and uses an adaptation motor (Box 2). Much attention has been paid to the molecular characterization of adaptation, as this approach should elicit strong candidates for members of the transduction apparatus.

Fast adaptation occurs on a millisecond to sub-millisecond timescale, and probably ensues when  $\text{Ca}^{2+}$  enters an open transduction channel, binds to a site near the channel, and causes the channel to close (Box 2). In slow adaptation, which requires tens to hundreds of milliseconds for completion, a motor molecule is hypothesized to



move up and down the actin core of a stereocilium to restore gating-spring tension towards its resting value<sup>60</sup>. Slow adaptation may be mediated by either myosin 1c (Myo1c or myosin I $\beta$ ) or by myosin VIIA (Myo7a) (Box 2).

### Essential molecules for transduction

Traditional biochemical and molecular-biological methods for identification of molecules essential for hair-cell function have been largely thwarted by hair cells' scarcity. The identification of 'deafness genes' by genetic screens has been the most productive approach used so far. Mutation of at least 100 human genes and 50 mouse genes lead to deafness associated with other dysfunctions (syndromic deafness) or deafness alone (nonsyndromic deafness).

But protein products of deafness genes generally do not have direct roles in mechanotransduction. Although unfortunate for those interested in transduction, this observation is unsurprising: the inner ear is a complicated structure, and sound detection relies on middle- and inner-ear formation, hair-bundle assembly, ionic homeostasis, synaptic transmission and a host of other events. Many of these deafness genes instead encode proteins required for development of the auditory or vestibular systems, for maintenance of the unusual extracellular fluid (endolymph), or for other essential functions not directly related to mechanical transduction<sup>61</sup>.

A few of the identified molecules do seem to carry out their critical roles in stereocilia. Molecules essential for hair bundles fall readily into several classes of molecules (described in Box 3): cytoskeletal components like actin, espin, Myo6, Myo7a, Myo15 and the mammalian homologue of diaphanous; cell-adhesion or junctional molecules like Cdh23, protocadherin-15 (Pcd15) and vezatin; and the calcium pump PMCA2. These molecules and others prominent in hair bundles like fimbrin and calmodulin can be assembled into a speculative and incomplete picture (Box 3). An important question is whether the already-identified molecules form a transduction complex or instead are important for other hair-bundle functions, such as maintaining structural integrity of the bundle. Loss of Myo7a, Pcd15 and Cdh23, at least, leads to disarrayed hair bundles, suggesting that these molecules help hold stereocilia together.

### Future directions for mechanotransduction research

Because of the difficulties in working with tiny, inaccessible touch cells, the genetic scheme of *C. elegans* touch transduction has yet to be confirmed by electrophysiological or biochemical experiments. Perhaps the model will be more easily approached in a vertebrate mechanoreceptive cell that is more amenable to such approaches; the relevance of transduction in such a cell type will depend on the molecular makeup of the transduction apparatus in various mechanoreceptive cells. Alternatively, the challenges of measuring mechanically sensitive currents in *C. elegans* may yet be overcome.

Characterization of mechanotransduction in *Drosophila* has just begun. Although there is strong evidence indicating that NompC contributes to *Drosophila* mechanotransduction, residual mechanically sensitive current remaining in the absence of NompC indicates that other transduction channels are present. In addition, little is known about the ultrastructure of the transduction apparatus — where exactly is it located and how do mechanical stimuli lead to channel opening? The combination of electrophysiology and powerful genetics afforded by the sequenced fly genome suggests that we can expect important developments over the coming few years.

A crucial question for the field is the generality of the results from invertebrate model systems. The touch-receptor transduction apparatus of *C. elegans* seems to be used by some vertebrate mechanoreceptors. Does conservation through the vertebrates also hold for the distinct *Drosophila* transduction apparatus? In addition to shared developmental-control molecules and intriguing structural similarities during development, *Drosophila* mechanotransduction exhibits several surprising physiological similarities to vertebrate transduction by hair cells. These similarities include a

high-K<sup>+</sup>-receptor endolymph made by supporting cells, directional selectivity, adaptation to sustained mechanical stimuli that is voltage- and size-dependent, microsecond response latencies, and sensitivity to nanometre-scale stimuli<sup>29</sup>. Although no obvious homologues of *nompC* are present in the human genome, the presence of dozens of orphan members of the broader TRP family suggests that a vertebrate transduction channel could still come from this family. Identification of more of the *remp* and *nomp* genes should assist in determining the generality of this transduction system.

Many questions remain about hair-cell transduction. The rough draft of the transduction apparatus comprising of the known proteins of the hair bundle is, to say the least, incomplete. Although two strong candidates have been advanced for the adaptation motor, we do not know how the motor assembles and how it interacts with other parts of the transduction apparatus. Although Cdh23 and Pcd15 each probably contribute to stereociliary crosslinks, which crosslinks contain these molecules remains unknown. Equally mysterious are the identities of two of the most interesting parts of the transduction apparatus, the transduction channel and tip link. Furthermore, the known human and mouse deafness genes do not constitute a saturating search of the genomes, and only a fraction of these have been cloned. Although the genetic approach is powerful, other important transduction molecules that are essential elsewhere in the organism will be missed. Screening for proteins that interact with essential hair-bundle proteins — such as those interacting with Myo7a — will continue to be a powerful method for discovering new bundle proteins, including those involved in transduction. The complete sequencing of the human and mouse genomes will no doubt spur more powerful approaches. The transduction channel and other important molecules are there — someone has probably already seen their sequences — but we still do not know how to recognize them. The next five years will undoubtedly be an exciting time, as the molecular clues that have begun to emerge will stimulate the right experiments to find the rest of the transduction apparatus. □

1. Chalfie, M. A molecular model for mechanosensation in *Caenorhabditis elegans*. *Biol. Bull.* **192**, 125–130 (1997).
2. Tavernarakis, N. & Driscoll, M. Molecular modeling of mechanotransduction in the nematode *Caenorhabditis elegans*. *Annu. Rev. Physiol.* **59**, 659–689 (1997).
3. Hudspeth, A. J. Hair-bundle mechanics and a model for mechanoelectrical transduction by hair cells. *Soc. Gen. Physiol. Ser.* **47**, 357–370 (1992).
4. Narins, P. M. & Lewis, E. R. The vertebrate ear as an exquisite seismic sensor. *J. Acoust. Soc. Am.* **76**, 1384–1387 (1984).
5. Sukharev, S. I., Blount, P., Martinac, B., Blattner, F. R. & Kung, C. A large-conductance mechanosensitive channel in *E. coli* encoded by *mscL* alone. *Nature* **368**, 265–268 (1994).
6. Kloda, A. & Martinac, B. Molecular identification of a mechanosensitive channel in archaea. *Biophys. J.* **80**, 229–240 (2001).
7. Chalfie, M. & Sulston, J. Developmental genetics of the mechanosensory neurons of *Caenorhabditis elegans*. *Dev. Biol.* **82**, 358–370 (1981).
8. Chalfie, M. & Au, M. Genetic control of differentiation of the *Caenorhabditis elegans* touch receptor neurons. *Science* **243**, 1027–1033 (1989).
9. Duggan, A., Ma, C. & Chalfie, M. Regulation of touch receptor differentiation by the *Caenorhabditis elegans* *mec-3* and *unc-86* genes. *Development* **125**, 4107–4119 (1998).
10. Way, J. C. & Chalfie, M. *mec-3*, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. *Cell* **54**, 5–16 (1988).
11. Savage, C. *et al.* *mec-7* is a beta-tubulin gene required for the production of 15-prototubulin microtubules in *Caenorhabditis elegans*. *Genes Dev.* **3**, 870–881 (1989).
12. Fukushige, T. *et al.* MEC-12, an  $\alpha$ -tubulin required for touch sensitivity in *C. elegans*. *J. Cell Sci.* **112**, 395–403 (1999).
13. Huang, M., Gu, G., Ferguson, E. L. & Chalfie, M. A stomatin-like protein necessary for mechanosensation in *C. elegans*. *Nature* **378**, 292–295 (1995).
14. Du, H., Gu, G., William, C. M. & Chalfie, M. Extracellular proteins needed for *C. elegans* mechanosensation. *Neuron* **16**, 183–194 (1996).
15. Gu, G., Caldwell, G. A. & Chalfie, M. Genetic interactions affecting touch sensitivity in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **93**, 6577–6582 (1996).
16. Alvarez de la Rosa, D., Canessa, C. M., Fyfe, G. K. & Zhang, P. Structure and regulation of amiloride-sensitive sodium channels. *Annu. Rev. Physiol.* **62**, 573–594 (2000).
17. Garcia-Anoveros, J., Ma, C. & Chalfie, M. Regulation of *Caenorhabditis elegans* degenerin proteins by a putative extracellular domain. *Curr. Biol.* **5**, 441–448 (1995).
18. Huang, M. & Chalfie, M. Gene interactions affecting mechanosensory transduction in *Caenorhabditis elegans*. *Nature* **367**, 467–470 (1994).
19. Garcia-Anoveros, J., Garcia, J. A., Liu, J. D. & Corey, D. P. The nematode degenerin UNC-105 forms ion channels that are activated by degeneration- or hypercontraction-causing mutations. *Neuron* **20**, 1231–1241 (1998).
20. Kizer, N., Guo, X. L. & Hruska, K. Reconstitution of stretch-activated cation channels by expression of the  $\alpha$ -subunit of the epithelial sodium channel cloned from osteoblasts. *Proc. Natl Acad. Sci. USA*

- 94, 1013–1018 (1997).
21. Adams, C. M. *et al.* Ripped pocket and pickpocket, novel *Drosophila* DEG/ENaC subunits expressed in early development and in mechanosensory neurons. *J. Cell Biol.* **140**, 143–152 (1998).
  22. Drummond, H. A., Price, M. P., Welsh, M. J. & Abboud, F. M. A molecular component of the arterial baroreceptor mechanotransducer. *Neuron* **21**, 1435–1441 (1998).
  23. Garcia-Anoveros, J., Samad, T. A., Woolf, C. J. & Corey, D. P. Transport and localization of the DEG/ENaC ion channel BNaC1α to peripheral mechanosensory terminals of dorsal root ganglia neurons. *J. Neurosci.* **21**, 2678–2686 (2001).
  24. Price, M. P. *et al.* The mammalian sodium channel BNC1 is required for normal touch sensation. *Nature* **407**, 1007–1011 (2000).
  25. Fricke, B. *et al.* Epithelial Na<sup>+</sup> channels and stomatin are expressed in rat trigeminal mechanosensory neurons. *Cell Tissue Res.* **299**, 327–334 (2000).
  26. Drummond, H. A., Abboud, F. M. & Welsh, M. J. Localization of beta and gamma subunits of ENaC in sensory nerve endings in the rat foot pad. *Brain Res.* **884**, 1–12 (2000).
  27. Hummler, E. & Horisberger, J. D. Genetic disorders of membrane transport. V. The epithelial sodium channel and its implication in human diseases. *Am. J. Physiol.* **276**, G567–G571 (1999).
  28. Kernan, M., Cowan, D. & Zuker, C. Genetic dissection of mechanosensory transduction: mechanoreception—defective mutations of *Drosophila*. *Neuron* **12**, 1195–1206 (1994).
  29. Walker, R. G., Willingham, A. T. & Zuker, C. S. A *Drosophila* mechanosensory transduction channel. *Science* **287**, 2229–2234 (2000).
  30. Eberl, D. F., Hardy, R. W. & Kernan, M. J. Genetically similar transduction mechanisms for touch and hearing in *Drosophila*. *J. Neurosci.* **20**, 5981–5988 (2000).
  31. Chung, Y. D., Zhu, J., Han, Y. & Kernan, M. J. *nompA* encodes a PNS-specific, ZP domain protein required to connect mechanosensory dendrites to sensory structures. *Neuron* **29**, 415–428 (2001).
  32. Legan, P. K., Rau, A., Keen, J. N. & Richardson, G. P. The mouse tectorins. Modular matrix proteins of the inner ear homologous to components of the sperm-egg adhesion system. *J. Biol. Chem.* **272**, 8791–8801 (1997).
  33. Sedgwick, S. G. & Smerdon, S. J. The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem. Sci.* **24**, 311–316 (1999).
  34. Rubtsov, A. M. & Lopina, O. D. Ankyrins. *FEBS Lett.* **482**, 1–5 (2000).
  35. Colbert, H. A., Smith, T. L. & Bargmann, C. I. OSM-9, a novel protein with structural similarity to channels, is required for olfaction, mechanosensation, and olfactory adaptation in *Caenorhabditis elegans*. *J. Neurosci.* **17**, 8259–8269 (1997).
  36. Liedtke, W. *et al.* Vanilloid receptor-related osmotically activated channel (VR-OAC), a candidate vertebrate osmoreceptor. *Cell* **103**, 525–535 (2000).
  37. Schweisguth, F., Gho, M. & Lecourtis, M. Control of cell fate choices by lateral signaling in the adult peripheral nervous system of *Drosophila melanogaster*. *Dev. Genet.* **18**, 28–39 (1996).
  38. Adam, J. *et al.* Cell fate choices and the expression of Notch, Delta and Serrate homologues in the chick inner ear: parallels with *Drosophila* sense-organ development. *Development* **125**, 4645–4654 (1998).
  39. Bermingham, N. A. *et al.* *Math1*: an essential gene for the generation of inner ear hair cells. *Science* **284**, 1837–1841 (1999).
  40. Ben-Arie, N. *et al.* Functional conservation of *atonal* and *Math1* in the CNS and PNS. *Development* **127**, 1039–1048 (2000).
  41. Corey, D. P. & Hudspeth, A. J. Kinetics of the receptor current in bullfrog saccular hair cells. *J. Neurosci.* **3**, 962–976 (1983).
  42. van Netten, S. M. & Kros, C. J. Gating energies and forces of the mammalian hair cell transducer channel and related hair bundle mechanics. *Proc. R. Soc. Lond. B* **267**, 1915–1923 (2000).
  43. Pickles, J. O., Comis, S. D. & Osborne, M. P. Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. *Hearing Res.* **15**, 103–112 (1984).
  44. Kachar, B., Parakkal, M., Kurc, M., Zhao, Y. & Gillespie, P. G. High-resolution structure of hair-cell tip links. *Proc. Natl Acad. Sci. USA* **97**, 13336–13341 (2000).
  45. Corey, D. P. & Hudspeth, A. J. Ionic basis of the receptor potential in a vertebrate hair cell. *Nature* **281**, 675–677 (1979).
  46. Crawford, A. C., Evans, M. G. & Fettiplace, R. The actions of calcium on the mechano-electrical transducer current of turtle hair cells. *J. Physiol.* **434**, 369–398 (1991).
  47. Ricci, A. J. & Fettiplace, R. Calcium permeation of the turtle hair cell mechanotransducer channel and its relation to the composition of endolymph. *J. Physiol.* **506**, 159–173 (1998).
  48. Kroese, A. B., Das, A. & Hudspeth, A. J. Blockage of the transduction channels of hair cells in the bullfrog's sacculus by aminoglycoside antibiotics. *Hear. Res.* **37**, 203–217 (1989).
  49. Rusch, A., Kros, C. J. & Richardson, G. P. Block by amiloride and its derivatives of mechano-electrical transduction in outer hair cells of mouse cochlear cultures. *J. Physiol.* **474**, 75–86 (1994).
  50. Sand, O. Effects of different ionic environments on the mechano-sensitivity of the lateral line organs in the mudpuppy. *J. Comp. Physiol. A* **102**, 27–42 (1975).
  51. Glowatzki, E., Ruppersberg, J. P., Zenner, H. P. & Rusch, A. Mechanically and ATP-induced currents of mouse outer hair cells are independent and differentially blocked by d-tubocurarine. *Neuropharmacology* **36**, 1269–1275 (1997).
  52. Baumann, M. & Roth, A. The Ca<sup>2+</sup> permeability of the apical membrane in neuromast hair cells. *J. Comp. Physiol.* **158**, 681–688 (1986).
  53. Jorgensen, F. & Kroese, A. B. Ca selectivity of the transduction channels in the hair cells of the frog sacculus. *Acta Physiol. Scand.* **155**, 363–376 (1995).
  54. Rusch, A. & Hummler, E. Mechano-electrical transduction in mice lacking the alpha-subunit of the epithelial sodium channel. *Hear. Res.* **131**, 170–176 (1999).
  55. Housley, G. D. *et al.* Expression of the P2X<sub>2</sub> receptor subunit of the ATP-gated ion channel in the cochlea: implications for sound transduction and auditory neurotransmission. *J. Neurosci.* **19**, 8377–8388 (1999).
  56. Housley, G. D., Luo, L. & Ryan, A. F. Localization of mRNA encoding the P2X<sub>2</sub> receptor subunit of the adenosine 5'-triphosphate-gated ion channel in the adult and developing rat inner ear by in situ hybridization. *J. Comp. Neurol.* **393**, 403–414 (1998).
  57. Kros, C. J., Rüsch, A. & Richardson, G. P. Mechano-electrical transducer currents in hair cells of the cultured neonatal mouse cochlea. *Proc. R. Soc. Lond. B* **249**, 185–193 (1992).
  58. Harteneck, C., Plant, T. D. & Schultz, G. From worm to man: three subfamilies of TRP channels. *Trends Neurosci.* **23**, 159–166 (2000).
  59. Hudspeth, A. J. & Gillespie, P. G. Pulling springs to tune transduction: adaptation by hair cells. *Neuron* **12**, 1–9 (1994).
  60. Howard, J. & Hudspeth, A. J. Mechanical relaxation of the hair bundle mediates adaptation in mechano-electrical transduction by the bullfrog's saccular hair cell. *Proc. Natl Acad. Sci. USA* **84**, 3064–3068 (1987).
  61. Steel, K. P. & Kros, C. J. A genetic approach to understanding auditory function. *Nature Genet.* **27**, 143–149 (2001).
  62. Liu, J., Schrank, B. & Waterston, R. H. Interaction between a putative mechanosensory membrane channel and a collagen. *Science* **273**, 361–364 (1996).
  63. Ricci, A. J. & Fettiplace, R. The effects of calcium buffering and cyclic AMP on mechano-electrical transduction in turtle auditory hair cells. *J. Physiol.* **501**, 111–124 (1997).
  64. Benser, M. E., Marquis, R. E. & Hudspeth, A. J. Rapid, active hair bundle movements in hair cells from the bullfrog's sacculus. *J. Neurosci.* **16**, 5629–5643 (1996).
  65. Ricci, A. J., Crawford, A. C. & Fettiplace, R. Active hair bundle motion linked to fast transducer adaptation in auditory hair cells. *J. Neurosci.* **20**, 7131–7142 (2000).
  66. Martin, P., Mehta, A. D. & Hudspeth, A. J. Negative hair-bundle stiffness betrays a mechanism for mechanical amplification by the hair cell. *Proc. Natl Acad. Sci. USA* **97**, 12026–12031 (2000).
  67. Gillespie, P. G. & Corey, D. P. Myosin and adaptation by hair cells. *Neuron* **19**, 955–958 (1997).
  68. Richardson, G. P. *et al.* A missense mutation in myosin VIIA prevents aminoglycoside accumulation in early postnatal cochlear hair cells. *Ann. NY Acad. Sci.* **884**, 110–124 (1999).
  69. Gillespie, P. G. & Hudspeth, A. J. High-purity isolation of bullfrog hair bundles and subcellular and topological localization of constituent proteins. *J. Cell Biol.* **112**, 625–640 (1991).
  70. Gibson, F. *et al.* A type VII myosin encoded by the mouse deafness gene *shaker-1*. *Nature* **374**, 62–64 (1995).
  71. Self, T. *et al.* Shaker-1 mutations reveal roles for myosin VIIA in both development and function of cochlear hair cells. *Development* **125**, 557–566 (1998).
  72. Ernest, S. *et al.* Mariner is defective in myosin VIIA: a zebrafish model for human hereditary deafness. *Hum. Mol. Genet.* **9**, 2189–2196 (2000).
  73. Hasson, T. *et al.* Unconventional myosins in inner-ear sensory epithelia. *J. Cell Biol.* **137**, 1287–1307 (1997).
  74. Bolz, H. *et al.* Mutation of *CDH23*, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D. *Nature Genet.* **27**, 108–112 (2001).
  75. Di Palma, F. *et al.* Mutations in *Cdh23*, encoding a new type of cadherin, cause stereocilia disorganization in waltzer, the mouse model for Usher syndrome type 1D. *Nature Genet.* **27**, 103–107 (2001).
  76. Ahmed, Z. M. *et al.* Mutations of the protocadherin gene *PCDH15* cause Usher syndrome type 1F. *Am. J. Hum. Genet.* **69**, 25–34 (2001).
  77. Alagramam, K. N. *et al.* Mutations in the novel protocadherin *PCDH15* cause Usher syndrome type 1F. *Hum. Mol. Genet.* **10**, 1709–1718 (2001).
  78. Alagramam, K. N. *et al.* The mouse Ames waltzer hearing-loss mutant is caused by mutation of *Pcdh15*, a novel protocadherin gene. *Nature Genet.* **27**, 99–102 (2001).
  79. Eudy, J. D. *et al.* Mutation of a gene encoding a protein with extracellular matrix motifs in Usher syndrome type 1a. *Science* **280**, 1753–1757 (1998).
  80. Kussel-Andermann, P. *et al.* Unconventional myosin VIIA is a novel A-kinase-anchoring protein. *J. Biol. Chem.* **275**, 29654–29659 (2000).
  81. Kussel-Andermann, P. *et al.* Vezatin, a novel transmembrane protein, bridges myosin VIIA to the cadherin-catenins complex. *EMBO J.* **19**, 6020–6029 (2000).
  82. Self, T. *et al.* Role of myosin VI in the differentiation of cochlear hair cells. *Dev. Biol.* **214**, 331–341 (1999).
  83. Lechler, T., Shevchenko, A. & Li, R. Direct involvement of yeast type I myosins in Cdc42-dependent actin polymerization. *J. Cell Biol.* **148**, 363–373 (2000).
  84. Zhao, Y., Yamoah, E. N. & Gillespie, P. G. Regeneration of broken tip links and restoration of mechanical transduction in hair cells. *Proc. Natl Acad. Sci. USA* **93**, 15469–15474 (1996).
  85. Kozel, P. J. *et al.* Balance and hearing deficits in mice with a null mutation in the gene encoding plasma membrane Ca<sup>2+</sup>-ATPase isoform 2. *J. Biol. Chem.* **273**, 18693–18696 (1998).
  86. Street, V. A., McKee-Johnson, J. W., Fonseca, R. C., Tempel, B. L. & Noben-Trauth, K. Mutations in a plasma membrane Ca<sup>2+</sup>-ATPase gene cause deafness in deafwaddler mice. *Nature Genet.* **19**, 390–394 (1998).
  87. Takahashi, K. & Kitamura, K. A point mutation in a plasma membrane Ca<sup>2+</sup>-ATPase gene causes deafness in Wriggle Mouse Sagami. *Biochem. Biophys. Res. Commun.* **261**, 773–778 (1999).
  88. Dumont, R. *et al.* Plasma membrane Ca<sup>2+</sup>-ATPase isoform 2a is the PMCA of hair bundles. *J. Neurosci.* **21**, 5066–5078 (2001).
  89. Yamoah, E. N. *et al.* Plasma membrane Ca<sup>2+</sup>-ATPase extrudes Ca<sup>2+</sup> from hair cell stereocilia. *J. Neurosci.* **18**, 610–624 (1998).
  90. Shepherd, G. M. G., Barres, B. A. & Corey, D. P. “Bundle-blot” purification and initial protein characterization of hair cell stereocilia. *Proc. Natl Acad. Sci. USA* **86**, 4973–4977 (1989).
  91. Walker, R. G., Hudspeth, A. J. & Gillespie, P. G. Calmodulin and calmodulin-binding proteins in hair bundles. *Proc. Natl Acad. Sci. USA* **90**, 2807–2811 (1993).
  92. Walker, R. G. & Hudspeth, A. J. Calmodulin controls adaptation of mechano-electrical transduction by hair cells of the bullfrog's sacculus. *Proc. Natl Acad. Sci. USA* **93**, 2203–2207 (1996).
  93. Noben-Trauth, K., Zheng, Q. Y., Johnson, K. R. & Nishina, P. M. *mdfw*: a deafness susceptibility locus that interacts with deaf waddler (*dfw*). *Genomics* **44**, 266–272 (1997).
  94. Flock, A., Cheung, H. C., Flock, B. & Utter, G. Three sets of actin filaments in sensory cells of the inner ear. Identification and functional orientation determined by gel electrophoresis, immunofluorescence/electron microscopy. *J. Neurocytol.* **10**, 133–147 (1981).
  95. Tilney, M. S. *et al.* Preliminary biochemical characterization of the stereocilia and cuticular plate of hair cells of the chick cochlea. *J. Cell Biol.* **109**, 1711–1723 (1989).
  96. Zheng, L. *et al.* The deaf jerker mouse has a mutation in the gene encoding the espin actin-bundling proteins of hair cell stereocilia and lacks espins. *Cell* **102**, 377–385 (2000).
  97. Lynch, E. D. *et al.* Nonsyndromic deafness DFNA1 associated with mutation of a human homolog of the *Drosophila* gene *diaphanous*. *Science* **278**, 1315–1318 (1997).

#### Acknowledgements

We thank M. Driscoll and N. Tavernarakis for the electron micrograph of Fig. 2 and T. A. Keil for the electron micrograph of Fig. 3.

# Molecular mechanisms of nociception

David Julius\* & Allan I. Basbaum†

\*Department of Cellular and Molecular Pharmacology, and †Departments of Anatomy and Physiology and W. M. Keck Foundation Center for Integrative Neuroscience, University of California San Francisco, San Francisco, California 94143, USA (e-mail: julius@socrates.ucsf.edu)

**The sensation of pain alerts us to real or impending injury and triggers appropriate protective responses. Unfortunately, pain often outlives its usefulness as a warning system and instead becomes chronic and debilitating. This transition to a chronic phase involves changes within the spinal cord and brain, but there is also remarkable modulation where pain messages are initiated — at the level of the primary sensory neuron. Efforts to determine how these neurons detect pain-producing stimuli of a thermal, mechanical or chemical nature have revealed new signalling mechanisms and brought us closer to understanding the molecular events that facilitate transitions from acute to persistent pain.**

**J**ust as beauty is not inherent in a visual image, pain is a complex experience that involves not only the transduction of noxious environmental stimuli, but also cognitive and emotional processing by the brain. Progress has been made in identifying cortical loci that process pain messages, but far greater advances have been made in understanding the molecular mechanisms whereby primary sensory neurons detect pain-producing stimuli, a process referred to as nociception. These insights have arisen predominantly from the analysis of sensory systems in mammals, as well as from studies of invertebrates. Of course, invertebrate organisms do not experience pain per se, but they do have transduction mechanisms that enable them to detect and avoid potentially harmful stimuli in their environment. These signalling pathways can be regarded as the evolutionary precursors of nociceptive processing in vertebrates, and genetic studies have facilitated the identification and functional characterization of molecules and signalling pathways that contribute to the detection of noxious stimuli in animals. Indeed, many of the receptors and ion channels we refer to here are related to molecules highlighted in the accompanying reviews on mechanosensation, vision or olfaction in flies or worms.

## The primary afferent nociceptor

Nearly a century ago, Sherrington proposed the existence of the nociceptor, a primary sensory neuron that is activated by stimuli capable of causing tissue damage<sup>1</sup>. According to this model, nociceptors have characteristic thresholds or sensitivities that distinguish them from other sensory nerve fibres. Electrophysiological studies have, in fact, shown the existence of primary sensory neurons that can be excited by noxious heat, intense pressure or irritant chemicals, but not by innocuous stimuli such as warming or light touch<sup>2</sup>. In this respect, acute pain can be regarded as a sensory modality much like vision or olfaction, where stimuli of a certain quality or intensity are detected by cells with appropriately tuned receptive properties.

## Many types of nociceptors for many types of pain

Pain is unique among sensory modalities in that electrophysiological recordings of single primary sensory fibres have been made in awake humans, allowing simultaneous measurement of psychophysical responses when regions of the head and body are stimulated<sup>3</sup>. Fibres that innervate regions of the head and body arise from cell bodies in trigeminal and dorsal

root ganglia (DRG), respectively, and can be categorized into three main groups based on anatomical and functional criteria (Fig. 1a). Cell bodies with the largest diameters give rise to myelinated, rapidly conducting A $\beta$  primary sensory fibres. Most, but not all<sup>4</sup>, A $\beta$  fibres detect innocuous stimuli applied to skin, muscle and joints and thus do not contribute to pain. Indeed, stimulation of large fibres can reduce pain, as occurs when you activate them by rubbing your hand. By contrast, small- and medium-diameter cell bodies give rise to most of the nociceptors, including unmyelinated, slowly conducting C fibres and thinly myelinated, more rapidly conducting A $\delta$  fibres, respectively. It has long been assumed that A $\delta$  and C nociceptors mediate 'first' and 'second' pain, respectively, namely the rapid, acute, sharp pain and the delayed, more diffuse, dull pain evoked by noxious stimuli<sup>5</sup> (Fig. 1b).

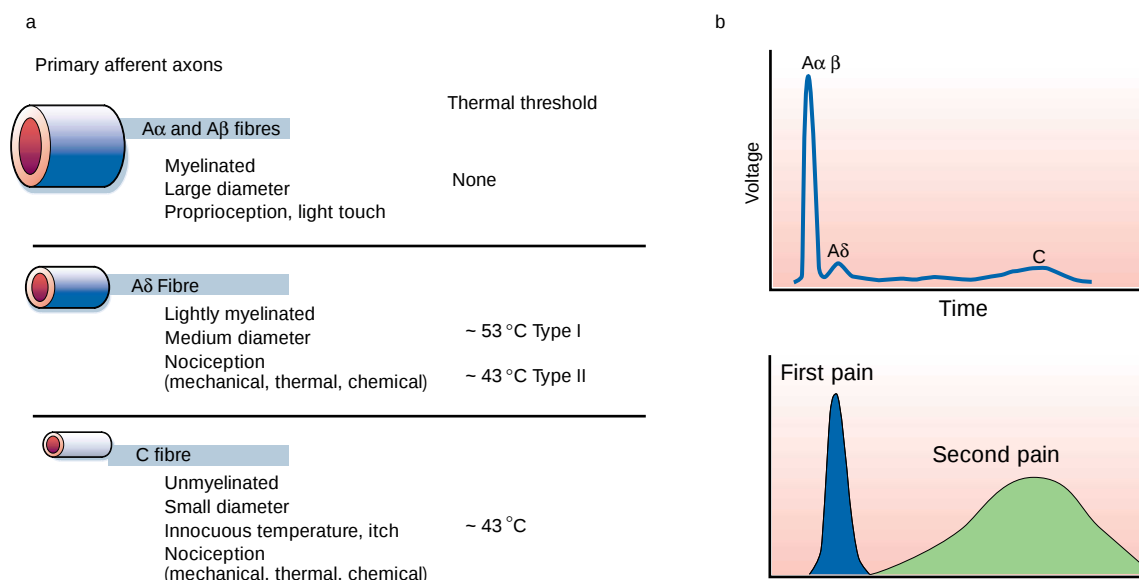
There are two main classes of A $\delta$  nociceptor<sup>6</sup>; both respond to intense mechanical stimuli, but can be distinguished by their differential responsiveness to intense heat or how they are affected by tissue injury. Most C-fibre nociceptors are also polymodal, responding to noxious thermal and mechanical stimuli<sup>6</sup>. Others are mechanically insensitive, but respond to noxious heat. Importantly, most C-fibre nociceptors also respond to noxious chemical stimuli, such as acid or capsaicin, the pungent ingredient in hot chilli peppers. Finally, the natural stimulus of some nociceptors is difficult to identify. These so-called 'silent' or 'sleeping' nociceptors are responsive only when sensitized by tissue injury<sup>7</sup>.

These nociceptor profiles derive largely from analysis of fibres that innervate skin. But very different features characterize nociceptors in other tissues<sup>6</sup>. For example, although corneal afferents can be activated by capsaicin and sensitized by inflammatory mediators, pain is normally produced by innocuous tactile stimulation. In teeth, almost any stimulus produces pain. Visceral pain is unique in that there are no first (fast) and second (slow) components; instead, pain is often poorly localized, deep and dull<sup>8</sup>. Tissue damage is also not required for visceral pain to occur; it can result from excessive distension (for example, of the colon). And the pain of ischaemia may have unique features that reflect innervation of vasculature by distinct subsets of acid-sensitive primary sensory nociceptors. These features illustrate the difficulty of defining a nociceptor based only on activation threshold or on whether its activation evokes pain.

## The neurochemistry of nociceptors

Glutamate is the predominant excitatory neurotransmitter in all nociceptors. Histochemical studies of adult DRG, however,





**Figure 1** Different nociceptors detect different types of pain. **a**, Peripheral nerves include small-diameter (A $\delta$ ) and medium- to large-diameter (A $\alpha$ ,  $\beta$ ) myelinated afferent fibres, as well as small-diameter unmyelinated afferent fibres (C). **b**, The fact that conduction velocity is directly related to fibre diameter is highlighted in the compound

action potential recording from a peripheral nerve. Most nociceptors are either A $\delta$  or C fibres, and their different conduction velocities (6–25 and ~1.0 m s<sup>-1</sup>, respectively) account for the first (fast) and second (slow) pain responses to injury. Panel **b** adapted from ref. 75.

reveal two broad classes of unmyelinated C fibre. The so-called peptidergic population contains the peptide neurotransmitter substance P, and expresses TrkA, the high-affinity tyrosine kinase receptor for nerve growth factor (NGF)<sup>9</sup>. The second population does not express substance P or TrkA, but can be labelled selectively with the  $\alpha$ -D-galactosyl-binding lectin IB<sub>4</sub>, and expresses P2X<sub>3</sub> receptors, a specific subtype of ATP-gated ion channel. This categorization is a first approximation at best — as additional molecular markers become available, new subsets are likely to be recognized. It is, however, unclear whether these neurochemically distinct groups represent different functional classes of nociceptor. Moreover, because expression of neurotransmitters, their receptors and other signalling molecules are dramatically altered after tissue or nerve injury<sup>10,11</sup>, both the significance and the complexity of nociceptor neurochemistry are increased.

### Diversity of nociceptor signalling

All sensory systems must convert environmental stimuli into electrochemical signals. In the case of vision or olfaction, primary sensory neurons need only detect one type of stimulus (light or chemical odors) and use redundant and convergent biochemical mechanisms to accomplish this goal (Fig. 2a). In this regard, nociception is unique because individual primary sensory neurons of the 'pain pathway' have the remarkable ability to detect a wide range of stimulus modalities, including those of a physical and chemical nature. Compared with sensory neurons of other systems, nociceptors must therefore be equipped with a diverse repertoire of transduction devices (Fig. 2b). At the same time, markedly different stimuli of a chemical (capsaicin and acid) or physical (heat) variety can excite nociceptors by activating a single receptor, enabling the cell to integrate information and respond to complex changes in the physiological environment.

Primary afferent nociceptors are also unique in the extent to which their receptive properties can be modulated. Thus, nociceptors not only signal acute pain, but also contribute to persistent and pathological pain conditions (allodynia) that occur in the setting of injury, wherein pain is produced by innocuous stimuli<sup>5,12</sup>. Allodynia can result from two different conditions: increased responsiveness of

spinal cord 'pain' transmission neurons (central sensitization), or lowering of nociceptor activation thresholds (peripheral sensitization). With central sensitization, pain can be produced by activity in non-nociceptive primary sensory fibres. Peripheral sensitization is produced when nociceptor terminals become exposed to products of tissue damage and inflammation, referred to collectively as the 'inflammatory soup' (Fig. 3). Such products include extracellular protons, arachidonic acid and other lipid metabolites, serotonin, bradykinin, nucleotides and NGF, all of which interact with receptors or ion channels on sensory nerve endings. Because nociceptors can release peptides and neurotransmitters (for example, substance P, calcitonin-gene-related peptide and ATP) from their peripheral terminals when activated by noxious stimuli, they are able to facilitate production of the inflammatory soup by promoting the release of factors from neighbouring non-neuronal cells and vascular tissue, a phenomenon known as neurogenic inflammation<sup>5</sup>.

In contrast to vision, olfaction or taste, sensory nerve endings that detect painful stimuli are not localized to a particular anatomical structure, but are instead dispersed over the body, innervating skin, muscle, joints and internal organs. Although this has made the biochemical analysis of nociceptive pathways particularly challenging, the combined application of electrophysiological, pharmacological and genetic methods are generating significant progress in understanding the molecular basis of nociceptor signalling. In some respects, the field can be compared to the state of cellular immunology some 25 years ago, when one of the main goals was to define biochemical markers for specific subsets of lymphocytes. The goal for nociceptor research is to elucidate signalling functions for key cell-surface markers and to assign physiological roles to molecularly defined sub-populations of sensory neurons.

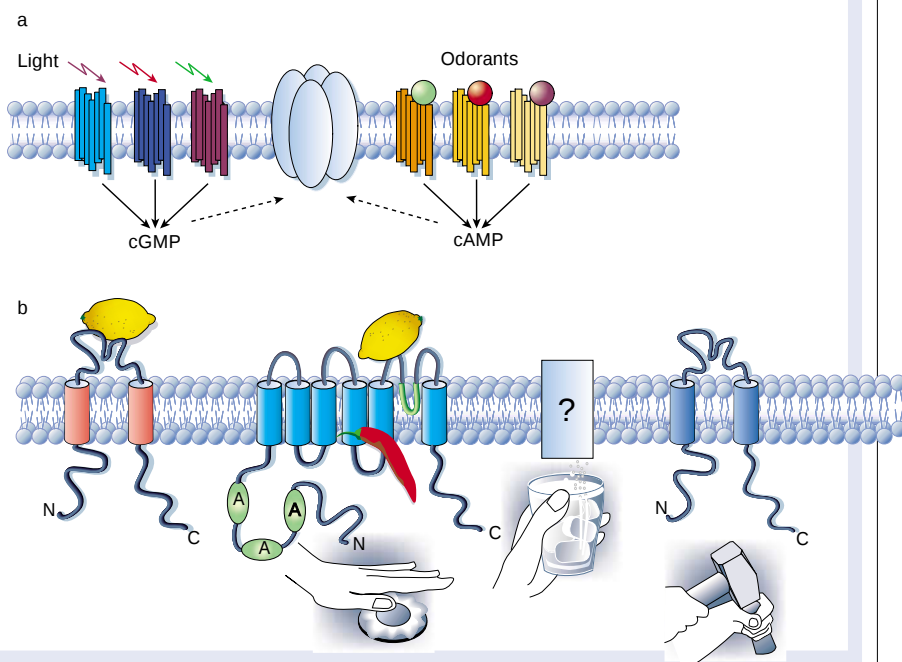
### Detectors of noxious stimuli

#### Response to heat

Remarkably, many functional characteristics of nociceptors are retained when sensory ganglia are dissociated and placed into culture<sup>13</sup>. Thus, ~45% of small- to medium-diameter neurons exhibit

**Figure 2** Polymodal nociceptors use a greater diversity of signal-transduction mechanisms to detect physiological stimuli than do primary sensory neurons in other systems.

**a**, In mammals, light or odorants are detected by a convergent signalling pathway in which G-protein-coupled receptors modulate the production of cyclic nucleotide second messengers, which then alter sensory neuron excitability by regulating the activity of a single type of cation channel. **b**, In contrast, nociceptors use different signal-transduction mechanisms to detect physical and chemical stimuli. Recent studies suggest that TRP-channel family members (VR1 and VRL-1) detect noxious heat, and that ENaC/DEG-channel family detect mechanical stimuli. Molecular transducers for noxious cold remain enigmatic. Noxious chemicals, such as capsaicin or acid (that is, extracellular protons) may be detected through a common transducer (VR1), illustrating aspects of redundancy in nociception. At the same time, a single type of stimulus can interact with multiple detectors, as shown by the ability of extracellular protons to activate not only VR1, but also ASICs, which are also members of the ENaC/DEG-channel family.



heat-evoked membrane currents with a 'moderate' threshold of  $\sim 45^{\circ}\text{C}$ , whereas another 5–10% of cells respond with a 'high' threshold of  $\sim 52^{\circ}\text{C}$  and are insensitive to capsaicin<sup>14,15</sup>. The former corresponds presumably to C and type II A $\delta$  nociceptors, and the latter to type I A $\delta$  nociceptors. What is the molecular basis by which specific thermal thresholds are established in these nociceptor subtypes? The answer came from a well-proven path of success in the pain field, namely identifying the molecular targets through which natural products (for example, morphine from the opium poppy or aspirin from willow bark) produce or modulate our sensation of pain.

In the case of moderate thermal nociception by C and type II A $\delta$  afferents, a transducer was revealed with the cloning and functional characterization of the vanilloid receptor VR1 (Fig. 2b), which is activated by capsaicin and other vanilloid compounds<sup>16</sup>. VR1 is a non-selective plasma-membrane cation channel possessing a very steep temperature dependence ( $Q_{10} = 20.6$ ) and a thermal activation threshold of  $\sim 43^{\circ}\text{C}$ , characteristics that are shared with native heat-evoked currents in sensory neurons<sup>17,18</sup>. The strong correlation between moderate heat and capsaicin sensitivity, and the similarity of the non-selective cationic currents and pharmacology underlying these responses<sup>14,15</sup> support the hypothesis that heat and capsaicin activate a common transducer. Heat-evoked single-channel currents are observed in membrane patches excised from sensory neurons or VR1-expressing cells, indicating that VR1 is an intrinsically heat-sensitive channel that functions as a molecular thermometer at the cell surface<sup>17</sup>.

The high correlation that exists between heat and capsaicin sensitivity in sensory neurons at the whole-cell level is less pronounced at the single-channel level<sup>19</sup>. Although this observation could be explained by different functional states of the channel, it has raised some controversy regarding the link between the activities of native and cloned channels. But studies of mice lacking functional VR1 channels<sup>20,21</sup> clearly show that cultured DRG neurons from VR1-null mice are severely deficient in moderate heat-evoked responses, whereas high-threshold heat responses persist. VR1<sup>-/-</sup> mice also have significantly fewer (>threefold) heat-sensitive C fibres, and total heat-evoked C-fibre output may be decreased by  $\sim 85\%$  compared with wild-type mice. Thus, although VR1 is not the only detector of moderate-threshold heat stimuli, it accounts for the majority of such responses and must contribute significantly to thermal coding in normal animals.

Somewhat paradoxically, VR1<sup>-/-</sup> mice showed normal behavioural responses at temperatures near the threshold for activation of VR1

and C fibres, but exhibited markedly reduced pain behaviour at higher temperatures ( $>50^{\circ}\text{C}$ ). In other words, VR1<sup>-/-</sup> animals recognize a noxious heat stimulus, but discriminate poorly among stimuli of different noxious intensities. Perhaps the threshold for a behavioural pain response is determined by the thermal threshold of a small number of nociceptors, whereas discrimination among suprathreshold temperatures requires information from a larger cohort of nociceptors that adequately encode stimulus intensity. In some respects, the phenotype of the VR1<sup>-/-</sup> mouse illustrates the futility of a long-lasting debate as to whether pain is generated by activity in specific nociceptors that are connected to the spinal cord through a dedicated pathway, or whether it is the pattern of activity across afferents that determines the pain response. Both models are likely to be relevant, depending on the magnitude of the stimulus and the context in which it is measured. Moreover, because most nociceptors are polymodal, our ability to distinguish pain sensations resulting from heat, cold or pressure must involve decoding of nociceptive signals within the central nervous system.

Another striking and significant phenotype of the VR1<sup>-/-</sup> mouse is its failure to develop increased sensitivity to heat in the context of tissue injury<sup>20,21</sup>. This deficit suggests that VR1 is modulated by one or more components of the inflammatory soup, which act on the nociceptor to increase its sensitivity to heat.

What mediates high-threshold heat responses in type I A $\delta$  afferents? One candidate transducer is the vanilloid receptor-like (VRL-1) channel (Fig. 2b) that shares  $\sim 50\%$  sequence identity with VR1 (ref. 22). Functional analysis of VRL-1 in transfected mammalian cells and frog oocytes showed that this channel is not responsive to vanilloid compounds, but can be activated by noxious thermal stimuli with a threshold of  $\sim 52^{\circ}\text{C}$ . VRL-1 is most prominently expressed by a subset of medium- to large-diameter myelinated neurons within the DRG. But because VRL-1 transcripts are also expressed outside the sensory nervous system, this channel probably serves multiple physiological functions<sup>22,23</sup>.

VR1 and VRL-1 belong to the larger family of transient receptor potential (TRP) channels, whose core transmembrane structure resembles that of voltage-gated potassium or cyclic nucleotide-gated channels (reviewed in ref. 24). The prototypical TRP channel was discovered in the *Drosophila* phototransduction pathway, where it is activated downstream of phospholipase C (PLC)-coupled rhodopsin (see review in this issue by Hardie and Raghu, pages 186–193). Some

mammalian TRP channels are also activated by G-protein-coupled or tyrosine kinase receptors that stimulate PLC, but as in the fly, the underlying gating mechanism remains enigmatic. Recent studies indicate that PLC-mediated hydrolysis of the membrane phospholipid phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P<sub>2</sub>) and the consequent production of lipid second messengers constitute important steps in TRP channel activation<sup>25</sup> (see below).

#### Detection of noxious cold

The definition of cold sensitivity is less stringent compared with heat sensitivity, in large part because thermal thresholds for activation of cold-sensitive fibres may not be as distinct or precipitous as they are with heat, and because the threshold for cold-evoked pain is not as precise. Whereas noxious heat (47 °C) activates ~50% of C fibres that innervate the hindpaw of a rodent, noxious cold (4 °C) may excite only 10–15% of fibres within the same cutaneous receptive field<sup>20</sup>. However, a significantly greater proportion of C and Aδ fibres are classified as cold-sensitive when the range of stimulus intensities extends below 0 °C (ref. 26). Few, if any, cold-sensitive (4 °C) afferents are heat sensitive, and VR1<sup>-/-</sup> mice show a normal prevalence of cold-responsive fibres in the hindpaw<sup>20</sup>, indicating that noxious heat and cold are detected by distinct mechanisms. It is not known whether noxious cold depolarizes nerve fibres by inhibiting a (Na<sup>+</sup> + K<sup>+</sup>)ATPase or background potassium current<sup>27</sup>, or by promoting calcium and/or sodium influx<sup>28,29</sup>.

#### Pressure and mechanical stress

Nociceptors can be activated by mechanical stress resulting from direct pressure, tissue deformation or changes in osmolarity, enabling the detection of touch, deep pressure, distension of a visceral organ, destruction of bone, or swelling. Functional, mechanically gated channels have yet to be identified at the molecular level in eukaryotes, although model genetic organisms such as bacteria, worms and flies have provided important leads for identifying mechanosensory transducers in mammals (see review in this issue by Gillespie and Walker, pages 194–202). One such ion channel of the degenerin (DEG/ENaC) family, called MDEG (alternatively, brain sodium channel 1 (BNC1) or acid-sensing ion channel 2 (ASIC2); Fig. 2b)<sup>30</sup>, has attracted particular interest in the pain field because its messenger RNA is expressed in primary sensory neurons<sup>31</sup>. Responses of primary sensory fibres from BNC1-deficient mice to a range of mechanical stimuli identified a specific class of rapidly adapting mechanoreceptors that showed reduced sensitivity to hair movement<sup>32</sup>. Other types of afferents, including C fibres, showed normal responses in these mutant mice, indicating that BNC1 is involved in some aspects of innocuous mechanical (touch) sensation, but not in the detection of noxious mechanical stimuli.

**Figure 3** The molecular complexity of the primary afferent nociceptor is illustrated by its response to inflammatory mediators released at the site of tissue injury. Some of the main components of the 'inflammatory soup' are shown, including peptides (bradykinin), lipids (prostaglandins), neurotransmitters (serotonin (5-HT) and ATP) and neurotrophins (NGF). The acidic nature of the inflammatory soup is also indicated. Each of these factors sensitize (lower the threshold) or excite the terminals of the nociceptor by interacting with cell-surface receptors expressed by these neurons. Examples of these factors and representative molecular targets are indicated in the box. Activation of the nociceptor not only transmits afferent messages to the spinal cord dorsal horn (and from there to the brain), but also initiates the process of neurogenic inflammation. This is an efferent function of the nociceptor whereby release of neurotransmitters, notably substance P and calcitonin gene related peptide (CGRP), from the peripheral terminal induces vasodilation and plasma extravasation (leakage of proteins and fluid from postcapillary venules), as well as activation of many non-neuronal cells, including mast cells and neutrophils. These cells in turn contribute additional elements to the inflammatory soup. Figure adapted from refs 75,76.

One way to detect pressure or tissue deformation is through activation of a mechanically gated protein. Another mechanism might involve a 'mechanochemical' process whereby stretch evokes the release of a diffusible chemical messenger that then excites nearby primary sensory nerve terminals. Extracellular ATP is of particular interest because large- and small-diameter sensory neurons express G-protein-coupled ATP receptors or ATP-gated ion channels (P2Y and P2X receptors, respectively), and because extracellular ATP excites primary sensory neurons<sup>33</sup>. Using frog oocytes as a model system, Nakamura and Strittmatter<sup>34</sup> showed that mechanical stimulation can release ATP from the cell, promoting autocrine activation of P2Y receptors on the cell's surface. *In vivo*, mechanical force might therefore promote ATP release from one or more cell types in the periphery, where it could activate purinergic receptors (for example, P2X<sub>3</sub>) on nearby nociceptor terminals.

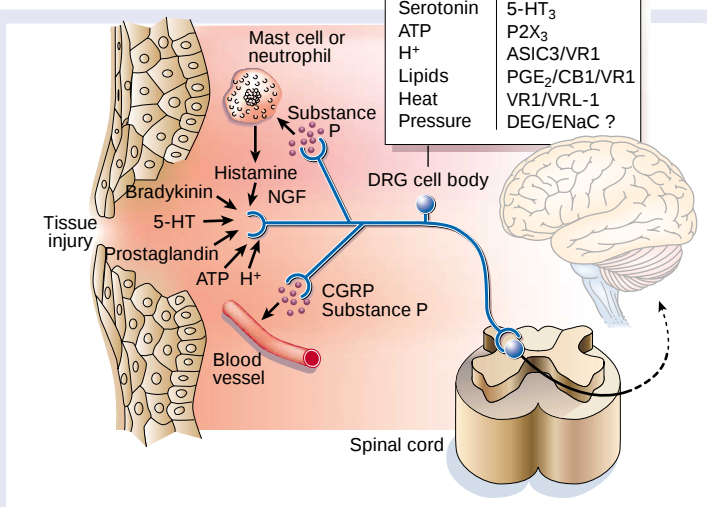
In fact, P2X<sub>3</sub>-deficient mice show reduced urination stemming from a failure to empty a full bladder<sup>35</sup>. Filling and stretching of the urinary bladder promotes the release of ATP from epithelial cells that lie beneath the smooth muscle layer. Once released, ATP may excite sensory nerve endings embedded within the bladder wall to initiate the voiding reflex. The deficit observed in P2X<sub>3</sub><sup>-/-</sup> mice provides compelling evidence that ATP transduces signals of mechanical tissue distension into depolarization of primary sensory neurons, in this case through direct activation of an ion channel.

#### Chemical transducers to make the pain worse

As described above, injury heightens our pain experience by increasing the sensitivity of nociceptors to both thermal and mechanical stimuli. This phenomenon results, in part, from the production and release of chemical mediators from the primary sensory terminal and from non-neuronal cells (for example, fibroblasts, mast cells, neutrophils and platelets) in the environment<sup>36</sup> (Fig. 3). Some components of the inflammatory soup (for example, protons, ATP, serotonin or lipids) can alter neuronal excitability directly by interacting with ion channels on the nociceptor surface, whereas others (for example, bradykinin and NGF) bind to metabotropic receptors and mediate their effects through second-messenger signalling cascades<sup>11</sup>. Considerable progress has been made in understanding the biochemical basis of such modulatory mechanisms.

#### Extracellular protons and tissue acidosis

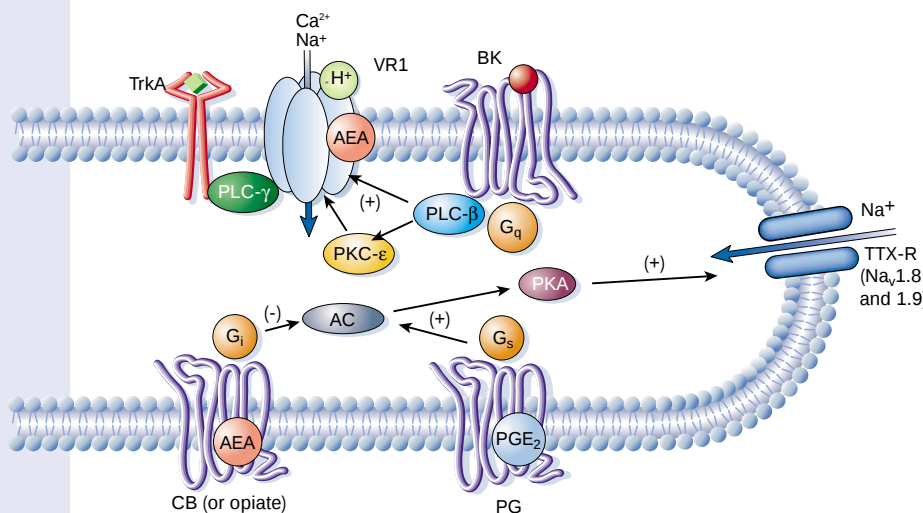
Local tissue acidosis is a hallmark physiological response to injury, and the degree of associated pain or discomfort is well correlated with the magnitude of acidification<sup>37</sup>. Application of acid (pH 5) to the skin produces sustained discharges in a third





**Figure 4** When nociceptors are exposed to products of injury and inflammation, their excitability is altered by a variety of intracellular signalling pathways. The figure highlights the vanilloid receptor (VR1) and tetrodotoxin-resistant (TTX-R) voltage-gated sodium channels (Na<sub>v</sub>1.8 and 1.9) as downstream targets of modulation. Responses of VR1 to heat can be potentiated by direct interaction of the channel with extracellular protons (H<sup>+</sup>) or lipid metabolites, such as anandamide (AEA). VR1 activity can also be heightened by agents such as NGF or bradykinin, which bind to their own cell-surface receptors (TrkA and BK, respectively) to stimulate phospholipase C (PLC-γ or PLC-β) signalling pathways. This, in turn, leads to hydrolysis of plasma membrane lipids and the subsequent stimulation of protein kinase C isoforms, such as PKC-ε. Both of these actions have been proposed to potentiate VR1 function.

Prostaglandins (PGE<sub>2</sub>) and other inflammatory products that activate adenyl cyclase (AC) through G<sub>s</sub>-coupled receptors also enhance nociceptor excitability. This occurs, in part, by a cyclic AMP-dependent protein kinase (PKA)-dependent phosphorylation of Na<sub>v</sub>1.8 and/or Na<sub>v</sub>1.9. By activating G<sub>i</sub>-coupled receptors, opiates and cannabinoids can counteract these increases in excitability of the nociceptor, and produce a peripherally mediated analgesia.



or more of polymodal nociceptors that innervate the receptive field<sup>20</sup>. At the cellular level, protons depolarize sensory neurons by directly activating a non-selective cationic current<sup>38,39</sup>. In many DRG neurons, this response consists of a transient, rapidly inactivating current that is carried predominantly by Na<sup>+</sup> ions, followed by a sustained non-selective cationic current. Responses of a sustained nature have been proposed to underlie persistent pain associated with tissue acidosis<sup>37</sup>, but this may not occur in all physiological settings. For example, during cardiac ischaemia, DRG neurons that innervate the epicardium show very large, but transient, response to extracellular protons<sup>40</sup>. A number of proton-sensitive channels are found on sensory neurons and thus an important goal has been to determine which, if any, of these molecules contributes to proton sensitivity of nociceptors *in vivo*. Two main candidates have emerged — VR1 and a family of ASICs.

The similarity between native proton (pH 5)-evoked and capsaicin-evoked currents in dissociated DRG neurons<sup>41</sup> is well established, and low extracellular pH can augment responses of cultured DRG neurons to capsaicin<sup>42</sup>. These observations suggested that protons and vanilloid compounds interact with the same ion-channel complex on the nociceptor (perhaps providing a cellular rationale for the culinary appeal of 'hot' and sour soup). Analysis of the cloned vanilloid receptor in heterologous expression systems has substantiated these observations. Protons have two main effects on VR1 function<sup>17</sup>. First, VR1 can be activated at room temperature when the extracellular pH drops below 6, producing currents that resemble the sustained component of proton-evoked responses observed in sensory neurons. Second, protons potentiate responses to capsaicin or heat, and do so over a concentration range (pH 6–8) that matches the extent of local acidosis associated with various forms of tissue injury. These changes in VR1 activity would be expected to increase nociceptor excitability, even at normal body temperature. Structure–function studies have now identified several negatively charged residues within putative extracellular loops of VR1 that are important for mediating these effects<sup>18,43</sup>, supporting the idea that protons interact directly with the vanilloid receptor to allosterically modulate channel function.

Although proton-evoked changes in VR1 thermal sensitivity closely resemble those exhibited by nociceptors during inflammation, does VR1 actually contribute to pH sensitivity of nociceptors *in vivo*? DRG neurons or sensory nerve fibres from VR1-deficient mice

do indeed show a marked reduction in sustained proton (pH 5)-evoked membrane currents. But proton-evoked responses, particularly those of a transient nature, are not eliminated completely in VR1<sup>−/−</sup> mice and may be mediated by members of the ASIC channel family.

Lazdunski and colleagues<sup>44</sup> described a family of two-transmembrane-domain proteins that are related to the putative mechanosensory DEG/ENaC channels. These novel cation-channel subunits were named ASICs because of their ability to be gated by reductions in extracellular pH when expressed in heterologous systems (Fig. 2b). Including splice variants, there are at least five ASIC subtypes (1a, 1b, 2a, 2b and 3) in rats, each having a unique profile of pH sensitivity, activation and desensitization rates, ionic permeability and tissue distribution. Most subtypes are expressed in DRG, with ASIC1b and ASIC3 (known also as ASIC-β and DRASIC, respectively) showing exclusive or preferential expression within sensory ganglia. Can ASIC channels account for any aspect of transient or sustained proton-evoked currents observed in DRG neurons? Heterologous expression of most ASIC subunits produces currents consisting of a single transient phase, or one having a sustained component that is Na<sup>+</sup>-selective or observed only at non-physiological proton concentrations (≤pH 5). However, co-expression of ASIC3 with ASIC2b (a splice variant of MDEG, also called MDEG2) generates a more native-like current containing a sustained component with non-selective cation permeability<sup>30</sup>. McCleskey and colleagues<sup>40</sup> have provided strong evidence that ASIC3/2b heteromeric channels, rather than VR1, underlie the unusually large and mostly transient proton-evoked currents observed in the relatively small subpopulation (2–3%) of DRG afferents that innervate the heart. This makes good physiological sense because occlusion of a cardiac artery produces modest acidosis (to just below pH 7), but this is sufficient to activate cardiac nociceptors or ASIC3/2b.

By contrast, genetic studies indicate that MDEG (BNC1) gene products (ASIC2a and 2b) do not contribute to acid sensitivity in most non-cardiac nociceptors<sup>32</sup>. Thus, BNC1<sup>−/−</sup> mice showed no obvious decrement in pH 5-evoked responses among large- or small-diameter DRG neurons. Moreover, acid produced normal sustained discharges in cutaneous C fibres from these animals. Gene knockouts of other members of this family will be needed to clarify the contribution of ASICs to proton detection and nociceptor sensitization at various physiological sites.

### Peptides and growth factors

Tissue damage promotes the release or production of bioactive peptides from non-neural cells and plasma proteins at the site of injury<sup>45</sup>. Principal among these is the nonapeptide bradykinin, which, when applied to primary sensory nerve terminals or cultured sensory neurons, produces immediate membrane depolarization as well as sensitization to other noxious, or even innocuous stimuli<sup>46</sup>. Bradykinin activates G-protein-coupled (BK<sub>2</sub>) receptors on these cells to stimulate PLC-catalysed hydrolysis of PtdIns(4,5)P<sub>2</sub>, consequently releasing Ca<sup>2+</sup> from intracellular stores and activating protein kinase C (PKC) (Fig. 4). Treatment of cultured DRG neurons with bradykinin augments heat-evoked currents in these cells<sup>47</sup>, an effect that may be mediated through direct or indirect modification of VR1 by the  $\epsilon$ -isoform of PKC<sup>48</sup> (Fig. 4). PKC- $\epsilon$ -knockout mice exhibit reduced thermal and mechanical hypersensitivity after treatment with adrenaline or acetic acid<sup>49</sup>, but effects of bradykinin have not been reported. In any event, molecular validation of this pathway will require biochemical proof that VR1 is phosphorylated in response to BK<sub>2</sub>-receptor activation and that mutation of one or more phosphate-accepting residues abrogates channel modulation.

Bradykinin might also heighten VR1 sensitivity through a PKC-independent process<sup>50</sup>. In this mechanism, channel modulation occurs as a direct consequence of PLC-mediated PtdIns(4,5)P<sub>2</sub> hydrolysis, in effect releasing VR1 from PtdIns(4,5)P<sub>2</sub>-mediated inhibition. Precedence for such a regulatory mechanism comes from studies of cyclic nucleotide-gated channels<sup>51</sup> and G-protein-gated inwardly rectifying potassium channels<sup>52,53</sup>. Moreover, as mentioned above, several members of the TRP channel family are activated downstream of PLC-coupled receptors, and recent evidence from both vertebrate and invertebrate systems suggests that PtdIns(4,5)P<sub>2</sub> hydrolysis also is important in modulating the activity of these channels<sup>24,25</sup>. Exogenously applied lipids, such as anandamide, arachidonate or diacylglycerol, have been shown to activate VR1 (see below) or other TRP channels (see review by Hardie and Raghu, pages 186–193), raising the possibility that these hydrophobic agonists exert their effects by displacing PtdIns(4,5)P<sub>2</sub> or other membrane lipids from an inhibitory site on the channel complex.

NGF is best known as a survival factor for embryonic neurons and is essential for the development of all primary sensory nociceptors. But in the adult, NGF serves a very different function, being released by mast cells, fibroblasts and other cell types at sites of injury and inflammation, where it acts on primary sensory nerve terminals to promote thermal hypersensitivity<sup>54</sup>. Consistent with this action, Mendell and colleagues<sup>55</sup> showed that exposure of cultured DRG neurons to NGF for several minutes produces acute sensitization of capsaicin-evoked responses. Although NGF elicits long-term changes in gene expression, its ability to promote short-term nociceptor sensitization suggests that post-translational mechanisms also are involved.

NGF binds to TrkA tyrosine kinase receptors on peptidergic sensory neurons to activate mitogen-activated protein (MAP) kinase and PLC- $\gamma$  signalling pathways<sup>56</sup>. When frog oocytes expressing TrkA and VR1 are treated with NGF for several minutes, proton-, vanilloid- or heat-evoked responses are potentiated, recapitulating the sensitization that occurs in cultured DRG neurons or primary sensory fibres. A TrkA mutant that specifically disrupts coupling of the receptor to PLC- $\gamma$  abrogates NGF enhancement of VR1, consistent with the potentiation mechanism outlined for bradykinin<sup>50</sup>. The importance of the MAP-kinase pathway in the regulation of gene expression is well appreciated, but the physiological significance of PLC activation to neurotrophin action has been enigmatic. These findings now suggest that PLC signalling contributes to NGF-mediated thermal hypersensitivity and possibly to other forms of short-term, post-translational nociceptor modulation (Fig. 4). VR1<sup>-/-</sup> mice do not develop thermal hypersensitivity in response to NGF or bradykinin treatment<sup>50</sup>, indicating that this channel is a probable downstream target for modulation by these and other

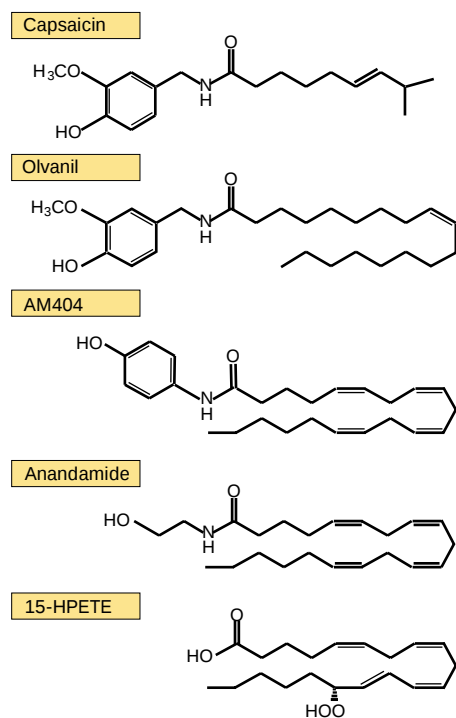
pro-algesic agents that activate PLC. TRP channels in the fly eye exist in a complex with other components of the phototransduction machinery (see review by Hardie and Raghu), an arrangement that influences sensitivity and kinetics of the signalling process<sup>57,58</sup>. Interestingly, VR1 forms a signalling complex with TrkA and PLC- $\gamma$  in heterologous systems<sup>50</sup>. Whether a similar signalling scaffold exists in nociceptors remains to be determined, but phylogenetic comparisons suggest that this might well be the case.

### Sensitization and activation by lipids

The efficacy of non-steroidal anti-inflammatory agents, such as aspirin, is attributed generally to blockade of cyclooxygenase (COX) enzymes that convert arachidonic acid, a lipid messenger, into pro-inflammatory prostanoic products, notably prostaglandin E<sub>2</sub> (PGE<sub>2</sub>). Most studies indicate that PGE<sub>2</sub> contributes to peripheral sensitization by binding to G-protein-coupled receptors that increase levels of cyclic AMP within nociceptors<sup>45</sup>. However, it now seems likely that cyclooxygenase products are also present in the spinal cord, where they could interact with receptors on the central terminals of nociceptors<sup>59</sup>. This idea has aroused great interest because it argues that COX inhibitors may exert their pain-relieving effects by modulating nociception at both peripheral and central sites<sup>60</sup>.

Recent studies have provided important information on a likely molecular target through which PGE<sub>2</sub> sensitizes primary sensory fibres. Nociceptors express a specific subclass of voltage-gated sodium channel that is resistant to blockade by tetrodotoxin<sup>61</sup>. These TTX-R Na<sup>+</sup> channels are believed to contribute significantly to action-potential firing rate and duration in small-diameter sensory neurons. Electrophysiological studies suggest that PGE<sub>2</sub> increases excitability of DRG neurons, in part by shifting the voltage dependence of TTX-R Na<sup>+</sup>-channel activation in the hyperpolarizing direction. This reduces the extent of membrane depolarization needed to initiate an action potential and favours repetitive spiking. Pharmacological and biochemical studies indicate that PGE<sub>2</sub>-dependent modulation of the TTX-R Na<sup>+</sup> current involves phosphorylation of the channel protein by cAMP-dependent protein kinase A<sup>62–64</sup> (Fig. 4). It is not known which of the two known TTX-R Na<sup>+</sup>-channel subtypes within DRG (Na<sub>v</sub>1.8 or Na<sub>v</sub>1.9)<sup>65–67</sup> are targeted by lipids, and under what conditions this occurs. Behavioural analysis of Na<sub>v</sub>1.8-deficient mice revealed modest deficits in acute sensation to noxious stimuli<sup>68</sup>. These animals also showed a delayed onset of inflammatory thermal hypersensitivity, but the maximal response was equivalent to that of wild-type animals. Such observations support the involvement of Na<sub>v</sub>1.8 in tissue injury-evoked hypersensitivity, but also suggest that there is redundancy among voltage-gated Na<sup>+</sup> channel subtypes in nociceptor function.

Capsaicin is a hydrophobic molecule that bears structural similarity to several lipid second messengers (Fig. 5), leading many to suggest that lipid metabolites might serve as endogenous vanilloid-receptor agonists. The hunt for such ligands has been further advanced by the realization that the vanilloid receptor belongs to the TRP channel family, some members of which can be activated by polyunsaturated fatty acids or other lipid metabolites (see review by Hardie and Raghu, pages 186–193). Indeed, we and others have shown that the endogenous cannabinoid-receptor agonist anandamide (arachidonyl ethanolamide), or lipoxygenase products of arachidonic-acid metabolism (for example, 12- and 15-(S)-hydroperoxyeicosatetraenoic acid), activate native or cloned vanilloid receptors when applied to whole cells or excised membrane patches containing VR1 channels<sup>69–71</sup>. These lipid messengers are admittedly weak as VR1 agonists (EC<sub>50</sub> ~ 1–10  $\mu$ M at 25 °C, pH 7.6), at least when compared to capsaicin, and this has fuelled some debate as to whether they should be considered 'endovanilloids'<sup>72</sup>. But the physiological setting in which these molecules are likely to act as VR1 agonists is one involving inflammation, and thus the relevant question should be whether they act synergistically with other



**Figure 5** An alignment of natural and synthetic vanilloid receptor agonists illustrates their structural similarity. Olvanil is a synthetic, non-pungent capsaicin analogue that activates VR1 with relatively slow kinetics. Anandamide is an endogenous lipid metabolite (similar in structure to arachidonic acid) that was initially discovered as a ligand for cannabinoid receptors. AM404 is a synthetic drug that blocks cellular re-uptake of anandamide. Both compounds activate native and cloned vanilloid receptors *in vitro* with relatively slow kinetics, similar to olvanil. 15-HPETE and other lipoxygenase products of arachidonic acid metabolism activate VR1 *in vitro* with potencies (1–10  $\mu$ M) resembling those of anandamide and AM404.

pro-inflammatory agents, such as bradykinin, NGF or protons, to facilitate VR1 gating and promote thermal hypersensitivity at sites of tissue injury. Finally, endogenous ligands with greater potency at VR1 may well exist, but these await discovery.

### The central terminal of the primary afferent nociceptor

All primary sensory nociceptors make synaptic connections with neurons in the grey matter (dorsal horn) of the spinal cord (Fig. 3). Subsets of dorsal horn neurons, in turn, project axons and transmit pain messages to higher brain centres, including the reticular formation, thalamus and ultimately the cerebral cortex<sup>5</sup>. Not surprisingly, the neural circuitry within the dorsal horn is incredibly complex, and much interest is focused on understanding whether subclasses of primary sensory nociceptors and the spinal circuits that they engage contribute differentially to the main classes of clinically relevant pains, namely, inflammatory pain resulting from tissue injury and neuropathic pain resulting from nerve injury<sup>9,12</sup>.

One of the most interesting questions concerning nociceptors is the extent to which receptors/transducers at the peripheral terminal are also functional at central (presynaptic) terminals in the spinal cord dorsal horn, and if so, whether they serve the same function. For some neurotransmitter receptors (for example, opioid receptors), the source of endogenous ligand is clear — interneurons in the superficial dorsal horn synthesize opioid peptides that can target presynaptic opioid receptors and regulate neurotransmitter release by decreasing  $Ca^{2+}$  conductance<sup>5</sup>. As noted above, it is likely that spinal cord-derived PGE<sub>2</sub> targets the central terminal of the

nociceptor. Functional presynaptic purinergic receptors have also been described and the source of ATP can be identified<sup>73</sup>. By contrast, presynaptic vanilloid receptors will never be exposed to noxious thermal temperatures or to the magnitude of pH changes that regulate VR1 gating at the peripheral terminal of the nociceptor, and thus other ligands must be considered. Lipid mediators, such as anandamide, may be relevant. There is evidence for presynaptic regulation of neurotransmitter release through an action of anandamide at both CB1 cannabinoid and vanilloid receptors in the dorsal horn<sup>74</sup>. Similarly, the function of central ASICs is unclear, nor is it known whether novel endogenous ligands for these channels exist.

### Future directions

One of the main challenges is to understand how both the specific physiological properties of nociceptors and the circuits that they engage in the central nervous system determine pain perception and resultant behaviour. Molecular markers make it possible to identify and manipulate the activity of subsets of nociceptors, so facilitating the mapping of spinal cord and brainstem circuits that are engaged by specific nociceptor populations. It is important to understand the function of the many neurotransmitters, receptors and transducers that are expressed by nociceptors and the significance of their transcriptional and post-translational regulation in the setting of injury. Although opioids and non-steroidal anti-inflammatory agents are analgesic drugs of choice for the treatment of pain, their utility is often limited by unacceptable side-effects due to actions at identical receptors outside of the pain pathway. Because many of the channels and receptors described in this review seem to be unique to the nociceptor (for example, TTX-R  $Na^+$  channels, P2X<sub>3</sub> and VR1), they represent promising targets for the development of new and highly selective local anaesthetics and analgesics for treating a wide variety of persistent pain conditions. □

- Sherrington, C. S. *The Integrative Action of the Nervous System* (Scribner, New York, 1906).
- Burgess, P. R. & Perl, E. R. Myelinated afferent fibres responding specifically to noxious stimulation of the skin. *J. Physiol.* **190**, 541–562 (1967).
- Weidner, C. et al. Functional attributes discriminating mechano-insensitive and mechano-responsive C nociceptors in human skin. *J. Neurosci.* **19**, 10184–10190 (1999).
- Djouhri, L., Bleazard, L. & Lawson, S. N. Association of somatic action potential shape with sensory receptive properties in guinea-pig dorsal root ganglion neurones. *J. Physiol.* **513**, 857–872 (1998).
- Basbaum, A. I. & Jessell, T. M. in *Principles of Neuroscience* (eds Kandel, E. R., Schwartz, J. H. & Jessell, T. M.) 472–491 (McGraw-Hill, New York, 2000).
- Raja, S. N., Meyer, R. A., Ringkamp, M. & Campbell, J. N. in *Textbook of Pain* (eds Wall, P. D. & Melzack, R.) 11–57 (Churchill Livingstone, Edinburgh, 1999).
- Schmidt, R. F. et al. Novel classes of responsive and unresponsive C nociceptors in human skin. *J. Neurosci.* **15**, 333–341 (1995).
- Gebhart, G. F. Visceral polymodal receptors. *Prog. Brain Res.* **113**, 101–112 (1996).
- Snider, W. D. & McMahon, S. B. Tackling pain at the source: new ideas about nociceptors. *Neuron* **20**, 629–632 (1998).
- Hökfelt, T., Zhang, X. & Wiesenfeld-Hallin, Z. Messenger plasticity in primary sensory neurons following axotomy and its functional implications. *Trends Neurosci.* **17**, 22–30 (1994).
- Woolf, C. J. & Salter, M. W. Neuronal plasticity: increasing the gain in pain. *Science* **288**, 1765–1769 (2000).
- Basbaum, A. I. & Woolf, C. J. Pain. *Curr. Biol.* **9**, R429–R431 (1999).
- Kress, M. & Reeh, P. W. More sensory competence for nociceptive neurons in culture. *Proc. Natl Acad. Sci. USA* **93**, 14995–14997 (1996).
- Nagy, I. & Rang, H. P. Noxious heat activates all capsaicin-sensitive and also a sub-population of capsaicin-insensitive dorsal root ganglion neurons. *Neuroscience* **88**, 995–997 (1999).
- Kirschstein, T., Greffrath, W., Busselberg, D. & Treede, R. D. Inhibition of rapid heat responses in nociceptive primary sensory neurons of rats by vanilloid receptor antagonists. *J. Neurophysiol.* **82**, 2853–2860 (1999).
- Caterina, M. J. et al. The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* **389**, 816–824 (1997).
- Tominaga, M. et al. The cloned capsaicin receptor integrates multiple pain-producing stimuli. *Neuron* **21**, 1–20 (1998).
- Welch, J. M., Simon, S. A. & Reinhart, P. H. The activation mechanism of rat vanilloid receptor 1 by capsaicin involves the pore domain and differs from the activation by either acid or heat. *Proc. Natl Acad. Sci. USA* **97**, 13889–13894 (2000).
- Nagy, I. & Rang, H. P. Similarities and differences between the responses of rat sensory neurons to noxious heat and capsaicin. *J. Neurosci.* **19**, 10647–10655 (1999).
- Caterina, M. J. et al. Impaired nociception and pain sensation in mice lacking the capsaicin receptor. *Science* **288**, 306–313 (2000).
- Davis, J. B. et al. Vanilloid receptor-1 is essential for inflammatory thermal hyperalgesia. *Nature* **405**, 183–187 (2000).
- Caterina, M. J., Rosen, T. A., Tominaga, M., Brake, A. J. & Julius, D. A capsaicin-receptor homologue with a high threshold for noxious heat. *Nature* **398**, 436–441 (1999).
- Kanzaki, M. et al. Translocation of a calcium-permeable cation channel induced by insulin-like



- growth factor-I. *Nature Cell Biol.* **1**, 165–170 (1999).
24. Harteneck, C., Plant, T. D. & Schultz, G. From worm to man: three subfamilies of TRP channels. *Trends Neurosci.* **23**, 159–166 (2000).
  25. Estacion, M., Sinkins, W. G. & Schilling, W. P. Regulation of *Drosophila* transient receptor potential-like (TrpL) channels by phospholipase C-dependent mechanisms. *J. Physiol.* **530**, 1–19 (2001).
  26. Simone, D. A. & Kajander, K. C. Responses of cutaneous A-fiber nociceptors to noxious cold. *J. Neurophysiol.* **77**, 2049–2060 (1997).
  27. Reid, G. & Flonta, M. Cold transduction by inhibition of a background potassium conductance in rat primary sensory neurones. *Neurosci. Lett.* **297**, 171–174 (2001).
  28. Askwith, C., Benson, C., Welsh, M. & Snyder, P. DEG/ENaC ion channels involved in sensory transduction are modulated by cold temperature. *Proc. Natl Acad. Sci. USA* **98**, 6459–6463 (2001).
  29. Suto, K. & Gotoh, H. Calcium signaling in cold cells studied in cultured dorsal root ganglion neurons. *Neuroscience* **92**, 1131–1135 (1999).
  30. Lingueglia, E. *et al.* A modulatory subunit of acid sensing ion channels in brain and dorsal root ganglion cells. *J. Biol. Chem.* **272**, 29778–29783 (1997).
  31. García-Añoveros, J., Samad, T. A., Zuvela-Jelaska, L., Woolf, C. J. & Corey, D. P. Transport and localization of the DEG/ENaC ion channel BNaC1 $\alpha$  to peripheral mechanosensory terminals of dorsal root ganglia neurons. *J. Neurosci.* **21**, 2678–2686 (2001).
  32. Price, M. P. *et al.* The mammalian sodium channel BNC1 is required for normal touch sensation. *Nature* **407**, 1007–1011 (2000).
  33. Krushal, O. A., Marchenko, S. M. & Obukhov, A. G. Cationic channels activated by extracellular ATP in rat sensory neurons. *Neuroscience* **27**, 995–1000 (1988).
  34. Nakamura, F. & Strittmatter, S. M. P2Y1 purinergic receptors in sensory neurons: contribution to touch-induced impulse generation. *Proc. Natl Acad. Sci. USA* **93**, 10465–10470 (1996).
  35. Cockayne, D. A. *et al.* Urinary bladder hyporeflexia and reduced pain-related behaviour in P2X<sub>2</sub>-deficient mice. *Nature* **407**, 1011–1015 (2000).
  36. Bevan, S. in *Textbook of Pain* (eds Wall, P. D. & Melzack, R.) 85–103 (Churchill-Livingstone, New York, 1999).
  37. Reeh, P. W. & Steen, K. H. Tissue acidosis in nociception and pain. *Prog. Brain Res.* **113**, 143–151 (1996).
  38. Krushal, O. A. & Pidoplichko, V. I. A receptor for protons in the nerve cell membrane. *Neuroscience* **5**, 2325–2327 (1980).
  39. Bevan, S. & Geppetti, P. Protons: small stimulants of capsaicin-sensitive sensory nerves. *Trends Neurosci.* **17**, 509–512 (1994).
  40. Sutherland, S. P., Benson, C. J., Adelman, J. P. & McCleskey, E. W. Acid-sensing ion channel 3 matches the acid-gated current in cardiac ischemia-sensing neurons. *Proc. Natl Acad. Sci. USA* **98**, 711–716 (2001).
  41. Bevan, S. & Yeats, J. Protons activate a cation conductance in a subpopulation of rat dorsal root ganglion neurones. *J. Physiol.* **433**, 145–161 (1991).
  42. Petersen, M. & LaMotte, R. H. Effect of protons on the inward current evoked by capsaicin in isolated dorsal root ganglion cells. *Pain* **54**, 37–42 (1993).
  43. Jordt, S. E., Tominaga, M. & Julius, D. Acid potentiation of the capsaicin receptor determined by a key extracellular site. *Proc. Natl Acad. Sci. USA* **97**, 8134–8139 (2000).
  44. Waldmann, R. *et al.* H<sup>+</sup>-gated cation channels. *Ann. NY Acad. Sci.* **868**, 67–76 (1999).
  45. Levine, J. D. & Reichling, D. B. in *Textbook of Pain* (eds Wall, P. D. & Melzack, R.) 59–84 (Churchill-Livingstone, New York, 1999).
  46. Burgess, G. M., Mullaney, I., McNeill, M., Dunn, P. M. & Rang, H. P. Second messengers involved in the mechanism of action of bradykinin in sensory neurons in culture. *J. Neurosci.* **9**, 3314–3325 (1989).
  47. Cesare, P. & McNaughton, P. A novel heat-activated current in nociceptive neurons and its sensitization by bradykinin. *Proc. Natl Acad. Sci. USA* **93**, 15435–15439 (1996).
  48. Premkumar, L. S. & Ahern, G. P. Induction of vanilloid receptor channel activity by protein kinase C. *Nature* **408**, 985–990 (2000).
  49. Khasar, S. G. *et al.* A novel nociceptor signaling pathway revealed in protein kinase C epsilon mutant mice. *Neuron* **24**, 253–260 (1999).
  50. Chuang, H. *et al.* Bradykinin and nerve growth factor release the capsaicin receptor from PtdIns(4,5)P<sub>2</sub>-mediated inhibition. *Nature* **411**, 957–962 (2001).
  51. Womack, K. B. *et al.* Do phosphatidylinositides modulate vertebrate phototransduction? *J. Neurosci.* **20**, 2792–2799 (2000).
  52. Huang, C. L., Feng, S. & Hilgemann, D. W. Direct activation of inward rectifier potassium channels by PIP<sub>2</sub> and its stabilization by G $\beta$  $\gamma$ . *Nature* **391**, 803–806 (1998).
  53. Zhang, H., He, C., Yan, X., Mirshahi, T. & Logothetis, D. E. Activation of inwardly rectifying K<sup>+</sup> channels by distinct PtdIns(4,5)P<sub>2</sub> interactions. *Nature Cell Biol.* **1**, 183–188 (1999).
  54. McMahon, S. B. & Bennett, D. L. H. in *Textbook of Pain* (eds Wall, P. D. & Melzack, R.) 105–128 (Harcourt, London, 1999).
  55. Shu, X. & Mendell, L. M. Nerve growth factor acutely sensitizes the response of adult rat sensory neurons to capsaicin. *Neurosci. Lett.* **274**, 159–162 (1999).
  56. Ganju, P., O'Bryan, J. P., Der, C., Winter, J. & James, I. F. Differential regulation of SHC proteins by nerve growth factor in sensory neurons and PC12 cells. *Eur. J. Neurosci.* **10**, 1995–2008 (1998).
  57. Chevesich, J., Kreuz, A. J. & Montell, C. Requirement for the PDZ domain protein, INAD, for localization of the TRP store-operated channel to a signaling complex. *Neuron* **18**, 95–105 (1997).
  58. Tsunoda, S. & Zuker, C. S. The organization of INAD-signaling complexes by a multivalent PDZ domain protein in *Drosophila* photoreceptor cells ensures sensitivity and speed of signaling. *Cell Calcium* **26**, 165–171 (1999).
  59. Malmberg, A. B. & Yaksh, T. L. Hyperalgesia mediated by spinal glutamate or substance P receptor blocked by spinal cyclooxygenase inhibition. *Science* **257**, 1276–1279 (1992).
  60. Samad, T. A. *et al.* Interleukin-1 $\beta$ -mediated induction of Cox-2 in the CNS contributes to inflammatory pain hypersensitivity. *Nature* **410**, 471–475 (2001).
  61. Baker, M. D. & Wood, J. N. Involvement of Na<sup>+</sup> channels in pain pathways. *Trends Pharmacol. Sci.* **22**, 27–31 (2001).
  62. England, S., Bevan, S. & Docherty, R. J. PGE<sub>2</sub> modulates the tetrodotoxin-resistant sodium current in neonatal rat dorsal root ganglion neurones via the cyclic AMP-protein kinase A cascade. *J. Physiol.* **495**, 429–440 (1996).
  63. Fitzgerald, E. M., Okuse, K., Wood, J. N., Dolphin, A. C. & Moss, S. J. cAMP-dependent phosphorylation of the tetrodotoxin-resistant voltage-dependent sodium channel SNS. *J. Physiol.* **516**, 433–446 (1999).
  64. Gold, M. S., Levine, J. D. & Correa, A. M. Modulation of TTX-R I<sub>Na</sub> by PKC and PKA and their role in PGE<sub>2</sub>-induced sensitization of rat sensory neurons *in vitro*. *J. Neurosci.* **18**, 10345–10355 (1998).
  65. Akopian, A. N., Silvilotti, L. & Wood, J. N. A tetrodotoxin-resistant voltage-gated sodium channel expressed by sensory neurons. *Nature* **379**, 257–262 (1996).
  66. Dib-Hajj, S. D., Tyrrell, L., Black, J. A. & Waxman, S. G. Na<sub>v</sub>1, a novel voltage-gated Na channel, is expressed preferentially in peripheral sensory neurons and downregulated after axotomy. *Proc. Natl Acad. Sci. USA* **95**, 8963–8968 (1998).
  67. Tate, S. *et al.* Two sodium channels contribute to the TTX-R sodium current in primary sensory neurons. *Nature Neurosci.* **1**, 653–655 (1998).
  68. Akopian, A. N. *et al.* The tetrodotoxin-resistant sodium channel SNS has a specialized function in pain pathways. *Nature Neurosci.* **2**, 541–548 (1999).
  69. Zygmunt, P. M. *et al.* Vanilloid receptors on sensory nerves mediate the vasodilator action of anandamide. *Nature* **400**, 452–457 (1999).
  70. Smart, D. *et al.* The endogenous lipid anandamide is a full agonist at the human vanilloid receptor (hVR1). *Br. J. Pharmacol.* **129**, 227–230 (2000).
  71. Hwang, S. W. *et al.* Direct activation of capsaicin receptors by products of lipoxygenases: endogenous capsaicin-like substances. *Proc. Natl Acad. Sci. USA* **97**, 6155–6160 (2000).
  72. Szolcsanyi, J. Anandamide and the question of its functional role for activation of capsaicin receptors. *Trends Pharmacol. Sci.* **21**, 203–204 (2000).
  73. Gu, J. G. & MacDermott, A. B. Activation of ATP P2X receptors elicits glutamate release from sensory neuron synapses. *Nature* **389**, 749–753 (1997).
  74. Tognetto, M. *et al.* Anandamide excites central terminals of dorsal root ganglion neurons via vanilloid receptor-1 activation. *J. Neurosci.* **21**, 1104–1109 (2001).
  75. Fields, H. L. *Pain* (McGraw-Hill, New York, 1987).
  76. Lembeck, F. & Gamse, R. Substance P in peripheral sensory processes. *Ciba Found. Symp.* **91**, 35–54 (1982).

## Acknowledgements

This work was supported by NIH, NIMH and an unrestricted grant from Bristol-Myers Squibb.

# How the olfactory system makes sense of scents

Stuart Firestein

Department of Biological Sciences, Columbia University, 923 Fairchild, MC 2438, New York 10027 USA (e-mail: sjf24@columbia.edu)

The human nose is often considered something of a luxury, but in the rest of the animal world, from bacteria to mammals, detecting chemicals in the environment has been critical to the successful organism. An indication of the importance of olfactory systems is the significant proportion — as much as 4% — of the genomes of many higher eukaryotes that is devoted to encoding the proteins of smell. Growing interest in the detection of diverse compounds at single-molecule levels has made the olfactory system an important system for biological modelling.

**T**he sensitivity and range of olfactory systems is remarkable, enabling organisms to detect and discriminate between thousands of low molecular mass, mostly organic compounds, which we commonly call odours. Represented in the olfactory repertoire are aliphatic and aromatic molecules with varied carbon backbones and diverse functional groups, including aldehydes, esters, ketones, alcohols, alkenes, carboxylic acids, amines, imines, thiols, halides, nitriles, sulphides and ethers. This remarkable chemical-detecting system, developed over eons of evolutionary time, has received considerable attention in the past decade, revealing sensing and signalling mechanisms common to other areas of the brain, but developed here to unusual sophistication.

How does the olfactory system manage this sophisticated discriminatory task? Beginning with the identification of a large family of G-protein-coupled receptors (GPCRs) in the nose, the foundations of a comprehensive understanding have emerged in surprisingly short order. The advent of advanced molecular and physiological techniques, as well as the publication of eukaryotic genomes from *Caenorhabditis elegans* to *Homo sapiens*, has provided the critical tools for unveiling some of the secrets. We now possess a detailed description of the transduction mechanism responsible for generating the stimulus-induced signal in primary sensory neurons, and also an explicit picture of the neural wiring, at least in the early parts of the system. From this body of work a view of molecular coding in the olfactory system has arisen that is surely incomplete, but nonetheless compelling in its simplicity and power.

Among higher eukaryotes, from flies through to mammals, there is a striking evolutionary convergence towards a conserved organization of signalling pathways in olfactory systems<sup>1</sup>. Two olfactory systems have developed in most animals. The common or main olfactory system is the sensor of the environment, the primary sense used by animals to find food, detect predators and prey, and mark territory. It is noteworthy for its breadth and discriminatory power. Like the immune complex, it is an open system built on the condition that it is not possible to predict, *a priori*, what molecules it (that is, you) might run into. Therefore, it is necessary to maintain an indeterminate but nonetheless precise sensory array. A second, or accessory, olfactory system has developed for the specific task of finding a receptive mate — a task of sufficient complexity that evolution has recognized the need for an independent and dedicated

system. Known as the vomeronasal system, it specializes in recognizing species-specific olfactory signals produced by one sex and perceived by the other, and which contain information not only about location but also reproductive state and availability. In addition to its role in sexual behaviours, it is important in influencing other social behaviours such as territoriality, aggression and suckling. This review will describe the recent advances that have emerged from molecular, physiological, imaging and genetic studies, and will highlight many of the remaining questions, especially as concerns the primary tasks of olfactory function: detecting, discriminating and signalling. Critical issues in areas such as development, gene regulation and higher processing, which are beyond the scope of this article, can be found in other recent reviews<sup>2,3</sup>.

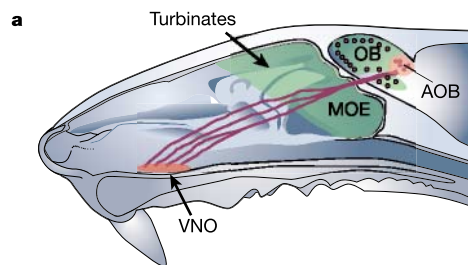
## Anatomical organization

### The sensory neuron

In vertebrates, the olfactory sensory neurons (OSNs) in the periphery are the primary sensing cell (Fig. 1b). Some 6–10 million of them form a neuroepithelium that lines a series of cartilaginous outcroppings, called turbinates, in the upper reaches of the nasal cavity in mammals; other vertebrates have similar specialized structures containing OSNs. The OSNs are bipolar neurons with a single dendrite that reaches up to the surface of the tissue and ends in a knob-like swelling from which project some 20–30 very fine cilia. These cilia, which actually lie in the thin layer of mucus covering the tissue, are the site of the sensory transduction apparatus. A thin axon from the proximal pole of the cell projects directly to higher brain regions (Fig. 1a,c). Invertebrates, particularly the arthropods, use a similar plan in which polarized neurons are specialized at one end for chemical detection and at the other for signalling.

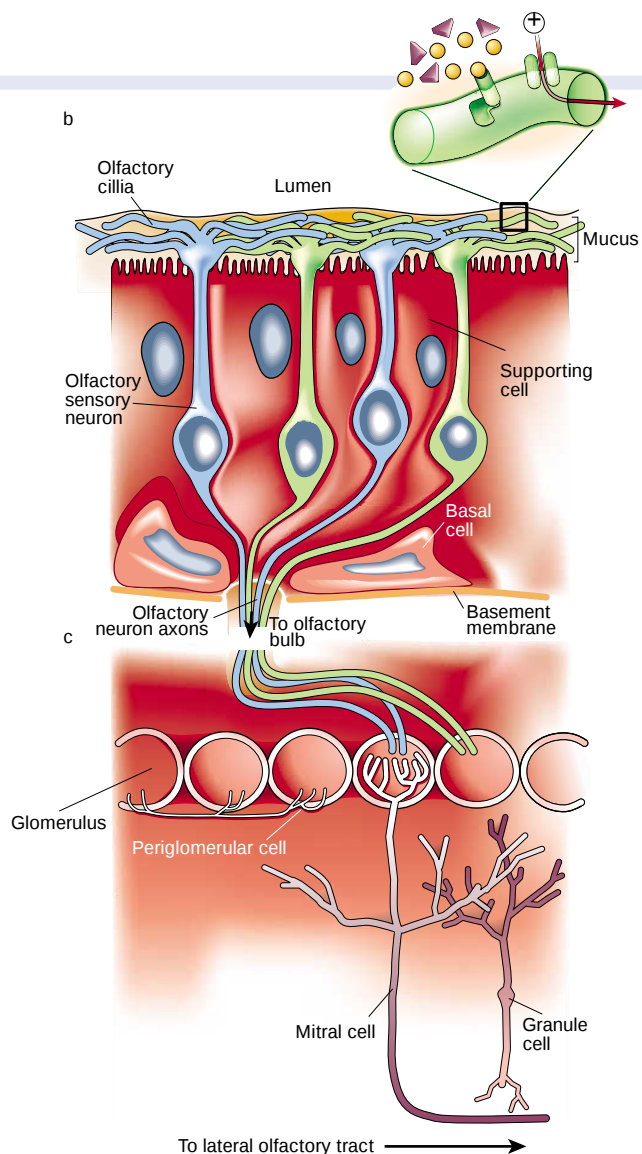
### Central pathways

OSNs send their axons into a region of the forebrain known as the olfactory bulb. Recent molecular-genetic studies using transgenic mice have shown that all the neurons expressing a particular receptor, no matter where they are found on the epithelial sheet, converge to a single 'target' in the olfactory bulb<sup>4</sup>. These targets are the glomeruli, spherical conglomerates of neuropil some 50–100  $\mu\text{m}$  in diameter that consist of the incoming axons of OSNs and the dendrites of the main projection cell in the bulb, the mitral cell (Fig. 1c). It is beyond the scope of this review to consider the complex wiring of the olfactory bulb, but at its simplest



**Figure 1** Functional anatomy and structure of the early olfactory system. In one of the clearest cases of function following form in the nervous system, the anatomy and structure of the early olfactory system reflect the strategy for discriminating between a large number of diverse stimuli. **a**, In a sagittal view of the rat head, the main olfactory epithelium (MOE) is highlighted in green. The turbinates are a set of cartilaginous flaps that serve to increase the surface area of the epithelium; they are covered with the thin olfactory neuroepithelium (shown in **b**). The cells of the MOE send their unbranched axons to targets in the olfactory bulb (OB) known as glomeruli (shown in **c**). The vomeronasal organ (VNO) is shown in red, and the targets of the VSN axons are in glomeruli in the accessory olfactory bulb (AOB). The structure of the nasal cavity is optimized for exposing the largest possible surface area of sensory neurons to a stimulus stream that is warmed, moistened and perhaps concentrated by sniffing.

**b**, The olfactory neuroepithelium is a relatively simple tissue consisting of only three cell types: olfactory sensory neurons (OSNs; the only neuronal cell type), supporting or sustentacular cells (a kind of glial cell, which possess microvilli on their apical surface), and a stem-cell population, known as basal cells, from which new OSNs are generated. **c**, Wiring of the early olfactory system. Each OSN expresses only one of the ~1,000 OR genes and the axons from all cells expressing that particular receptor converge onto one or a few 'glomeruli' in the OB. The nearly 2,000 glomeruli in the rat OB are spherical knots of neuropil, about 50–100  $\mu\text{m}$  in diameter, which contain the incoming axons of OSNs and the apical dendrites of the main input-output neuron of the OB, the mitral cell. Mitral axons leaving the OB project widely to higher brain structures including the piriform cortex, hippocampus and amygdala. Lateral processing of the message occurs through two populations of interneurons: periglomerular cells and granule cells.



the mitral cells receive information from the OSNs, and their axons in turn project to various higher brain regions. In one of the most extreme cases of convergence in the nervous system, several thousand OSN axons synapse on to the dendrites of only 5–25 mitral cells in each glomerulus.

#### Odour receptors in vertebrates and mammals

The discovery and publication a decade ago of the mammalian family of odour receptors<sup>5</sup> produced one anticipated and one surprising result. Expected was that olfactory receptors (ORs) are GPCRs similar to those known to be important in neurotransmission, photoreception (rhodopsin is a GPCR) and many other cellular processes. Unanticipated was the astonishing finding that there are as many as 1,000 genes for ORs in the mammalian genome, making it by far the largest family of GPCRs, and probably the largest gene family in the entire genome (Fig. 2). Fish and amphibians are less well endowed with about 100 ORs, and in the human genome an unexplained 60% of OR genes seem to be pseudogenes<sup>6</sup>.

Vertebrate odour receptors share many features with other GPCRs, including a coding region that lacks introns, a structure that predicts seven  $\alpha$ -helical membrane-spanning domains connected by intracellular and extracellular loops of variable lengths, and numerous conserved short sequences. But there are certain characteristics specific to ORs, such as an unusually long second extracellular loop, an extra pair of conserved cysteines in that loop, and other short

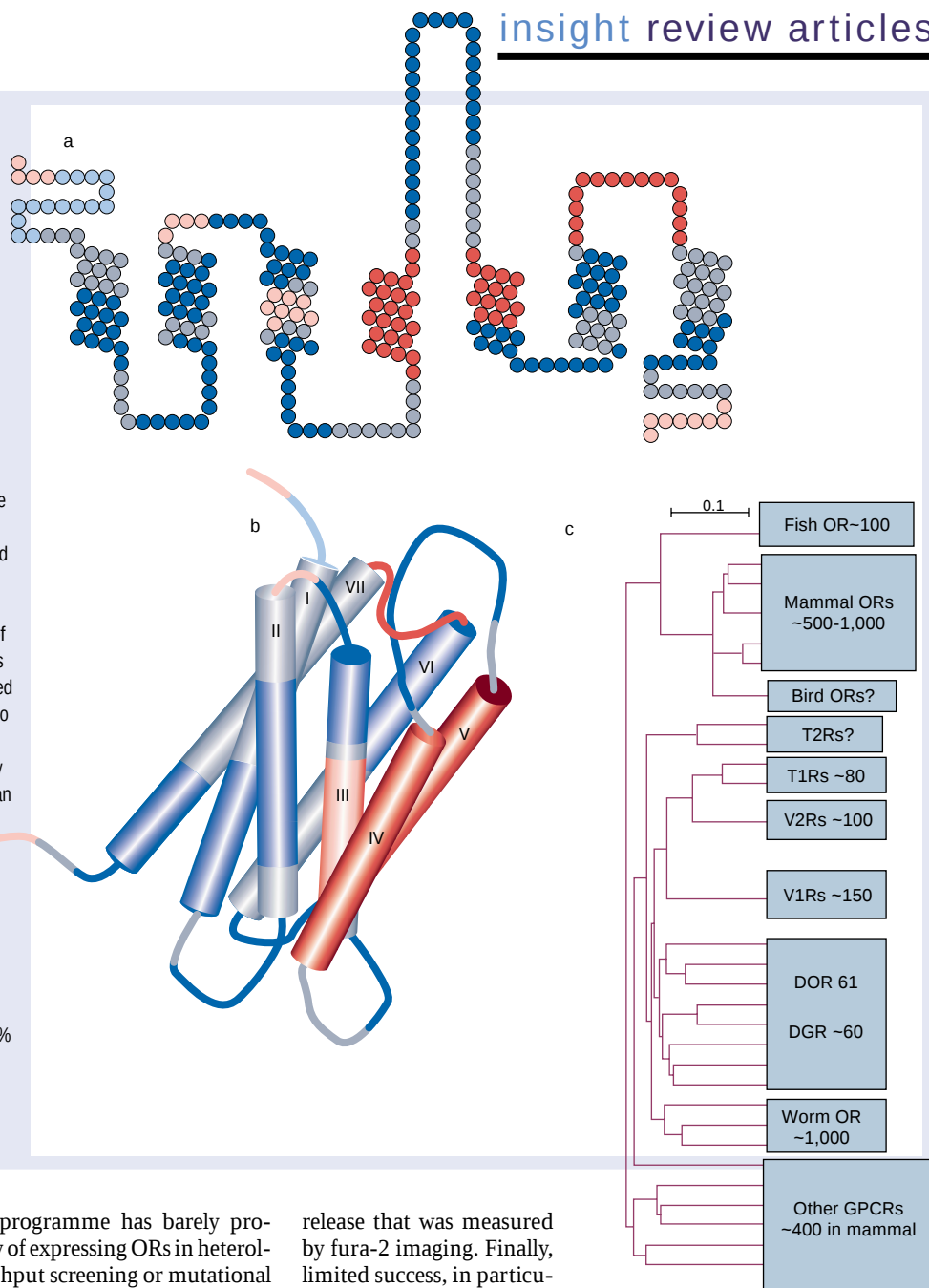
sequences (see a recent review by Mombaerts<sup>3</sup> for a detailed analysis of the receptor gene). Within the family of ORs there is a range of similarity, from less than 40% to over 90% identity. Perhaps most interesting is that there is a region of hypervariability, where the sequences show particularly strong divergence, in the third, fourth and fifth transmembrane regions. In three-dimensional models of the GPCRs, these three  $\alpha$ -helical barrels are thought to face each other and form a pocket about one-third of the way into the membrane<sup>7</sup> (Fig. 2). Based on studies from other GPCRs of this class (for example adrenergic receptors<sup>8</sup>), and in common with the binding site of retinal in rhodopsin<sup>9</sup>, this pocket is the probable binding site for ligands. The variability observed among the ORs in this region provides the first molecular basis for understanding the range, diversity and large number of olfactory ligands that can be detected and discriminated.

Such features make these receptors excellent models for structure–function studies in this class of receptors. In the case of a receptor known as I7, for example, the mouse and rat orthologues showed a differential response with one being more sensitive to octanal and the other to heptanal. Among the 15 amino acids that are different in the two genes, a single residue in transmembrane domain 5 (valine or isoleucine) was found to be sufficient to confer this alternate ligand sensitivity<sup>10</sup>. These sorts of results provide the impetus for developing a pharmacology of odour receptors that would produce activity matrices of large numbers of receptors tested against equally large chemical libraries.



**Figure 2** Odorant receptors are the jewel of olfactory research in the past 10 years.

The odorant receptors comprise the largest family of GPCRs. In mammals, odour receptors are represented by as many as 1,000 genes and may account for as much as 2% of the genome. Sequence comparison across the receptors has revealed many regions of conservation and variability that may be related to function. **a**, In a 'snake' diagram showing the amino acids for a particular receptor (M71), those residues that are most highly conserved are shown in shades of blue and those that are most variable are shown in shades of red. The seven  $\alpha$ -helical regions (boxed) are connected by intracellular and extracellular loops. **b**, A schematic view of the proposed three-dimensional structure of the receptor based on the recently solved structure of rhodopsin. Each of the transmembrane regions is numbered according to that model. The conserved (blue) and variable (red) regions are sketched onto this qualitative view and suggest that a ligand-binding region may be at least partially formed by the variable regions of the receptor. **c**, Mammalian odour receptors are related phylogenetically to other chemosensory receptors. In the tree depicted here the numbers refer to the approximate number of receptors in each family. OR, Odorant receptors; T1R, T2R, taste receptors; V3R, vomeronasal receptors; DOR, DGR, *Drosophila* odour and gustatory receptors; worm refers to *C. elegans*. The scale bar is a graphical distance equal to 10% sequence divergence.



But this ambitious experimental programme has barely progressed, owing to the puzzling difficulty of expressing ORs in heterologous systems suitable for high-throughput screening or mutational analysis. The main difficulty seems to be one of protein trafficking. For unknown reasons, the OR protein, although produced in transfected cells, seems to be trapped in endoplasmic reticulum, Golgi and endosomal compartments, with little or no receptor finding its way to the membrane<sup>11</sup>. In one solution to this problem, a recombinant receptor was engineered into an adenovirus, which was then used to infect OSNs in rat olfactory epithelia<sup>12</sup>. The infected OSNs were driven to express the recombinant receptor (the I7 receptor was used in this case) so that infected epithelia produced larger physiological responses to the ligands for that receptor than to other odours. In this way the first positive identification of a family of ligands was determined for a single OR. The ability of OSNs to express cloned receptors while other cells could not is further evidence for the involvement of some olfactory-specific chaperone or co-factor necessary for functional receptor expression. What that co-factor might be remains the focus of significant research efforts.

A few other ORs have been expressed in heterologous cell systems by including additional amino acids on the amino terminal. Fusing 20 amino acids from rhodopsin<sup>10</sup> or a short piece of serotonin receptor<sup>13</sup> to the N terminal of several ORs has enabled low levels of membrane expression in HEK 293 cells. Paired with expression of the promiscuous G protein ( $G_{\alpha 15}$ ), odours induced intracellular calcium

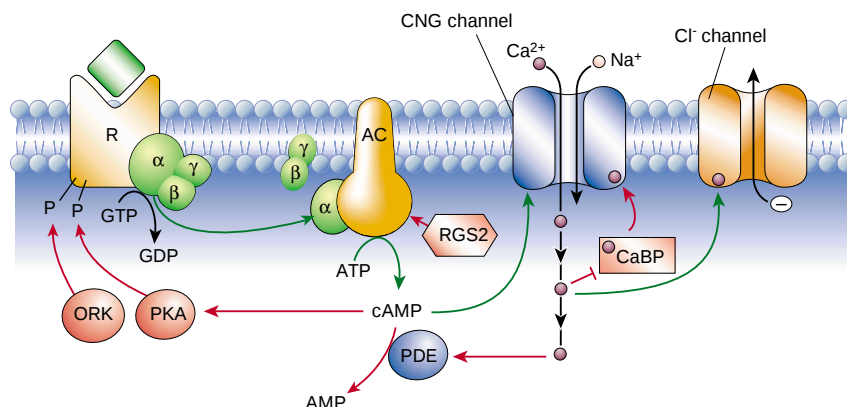
release that was measured by fura-2 imaging. Finally, limited success, in particular with ORs from fish, has been achieved in oocytes<sup>14</sup>. But in general there remains no robust, reliable, efficient system for expressing and assaying ORs. The eventual solution of this problem will surely initiate a research bonanza in GPCR receptor–ligand and structure–activity research. Molecular sensing by GPCRs, so well developed in the olfactory sense, is biologically ubiquitous, and over 50% of the pharmaceuticals currently on the market or in development are targeted at these receptors.

#### Invertebrate odour receptors

In what is something of a reversal of the standard strategy, mammalian molecular studies have often preceded those in invertebrates. The first invertebrate family of ORs were identified in *C. elegans*, but they are unrelated to those of any other species investigated<sup>15</sup>. They comprise an independent super family of 500+ ORs. The nematode olfactory system seems to be built on a strategy more comparable to vertebrate taste than olfaction, with multiple receptors expressed in a few sensory neurons. Nonetheless, many important accessory molecules have been discovered in this system and future genomic and structural research seems promising.

The appearance of the *Drosophila* odour receptor (DOR) family was perhaps the most important advance in olfactory studies in the

**Figure 3** Sensory transduction. Within the compact cilia of the OSNs a cascade of enzymatic activity transduces the binding of an odour molecule to a receptor into an electrical signal that can be transmitted to the brain. As described in detail in the text, this is a classic cyclic nucleotide transduction pathway in which all of the proteins involved have been identified, cloned, expressed and characterized. Additionally, many of them have been genetically deleted from strains of mice, making this one of the most investigated and best understood second-messenger pathways in the brain. AC, adenylyl cyclase; CNG channel, cyclic nucleotide-gated channel; PDE, phosphodiesterase; PKA, protein kinase A; ORK, olfactory receptor kinase; RGS, regulator of G proteins (but here acts on the AC); CaBP, calmodulin-binding protein. Green arrows indicate stimulatory pathways; red indicates inhibitory (feedback).



past five years<sup>16,17</sup>. These elusive receptors number at least 60 in the adult fly, with additional chemosensory receptors in the larva. They bear no homology to vertebrate ORs, and indeed have very little similarity with each other. Their classification as a family relies on a mildly conserved region in the seventh transmembrane domain and their common expression in olfactory tissues. They do share, with the vertebrate ORs, the unfortunate attribute of not expressing functionally in heterologous expression systems. Thus no DOR has been paired definitively with a cognate ligand. However, it seems that the *Drosophila* system is organized along the lines of the vertebrate system<sup>18</sup>, with each sensory neuron expressing only a single OR (with one curious exception, a single receptor that is also expressed in nearly every OSN), and all cells expressing the same receptor contacting a single glomerulus in the antennal lobe (a structure analogous to the olfactory bulb).

Because the fly is structurally more stereotypical, it is possible to map odour sensitivity across an antenna, and there are patterns of sensitivity that suggest a tight control on how receptors are expressed. With its similarity to the vertebrate system, its numerically simpler receptor repertoire and its genetic tractability, the *Drosophila* olfactory system should be very useful for investigating crucial issues of gene regulation, axon targeting and stimulus coding. And the potential value of utilizing insect olfaction as part of an integrated pest management strategy should not be overlooked. With the fly as a model, identification of receptors in insects that have prominent roles in agriculture and public health may lead to the discovery of repellents and attractants that can alter insect behaviour, without the disagreeable side effects of neurotoxic insecticides.

### Signal transduction

Once the receptor has bound an odour molecule, a cascade of events is initiated that transforms the chemical energy of binding into a neural signal (that is, a change in the membrane potential of the OSN). Although still obscure in invertebrates, this process is now generally well understood in mammals and other vertebrates (Fig. 3).

The ligand-bound receptor activates a G protein (an olfactory-specific subtype,  $G_{olf}$ ), which in turn activates an adenylyl cyclase (ACIII). The cyclase converts the abundant intracellular molecule ATP into cyclic AMP, a molecule that has numerous signalling roles in cells. In the case of OSNs the cAMP binds to the intracellular face of an ion channel (a cyclic nucleotide-gated (CNG) channel closely related to that found in photoreceptors; see review in this issue by Hardie and Raghu, pages 186–193), enabling it to conduct cations such as  $Na^+$  and  $Ca^{2+}$  (ref. 19). Inactive OSNs normally maintain a resting voltage across their plasma membrane of about  $-65$  mV (inside with respect to outside). When the CNG channels open, the influx of  $Na^+$  and  $Ca^{2+}$  ions causes the inside of the cell to become less negative. If enough channels are open for long enough, causing the

membrane potential to become about 20 mV less negative, the cell reaches threshold and generates an action potential. The action potential is then propagated along the axon, which crosses through a thin bone known as the cribriform plate, and into the forebrain where it synapses with second-order neurons in the olfactory bulb. Genetically altered mice in which various components of this transduction cascade have been deleted ( $G_{olf}$ , ACIII and most notably the CNG channel<sup>20–22</sup>) indicate that the cAMP pathway is the common pathway for all OSNs; involvement of other second messengers in modulatory roles awaits conclusive proof.

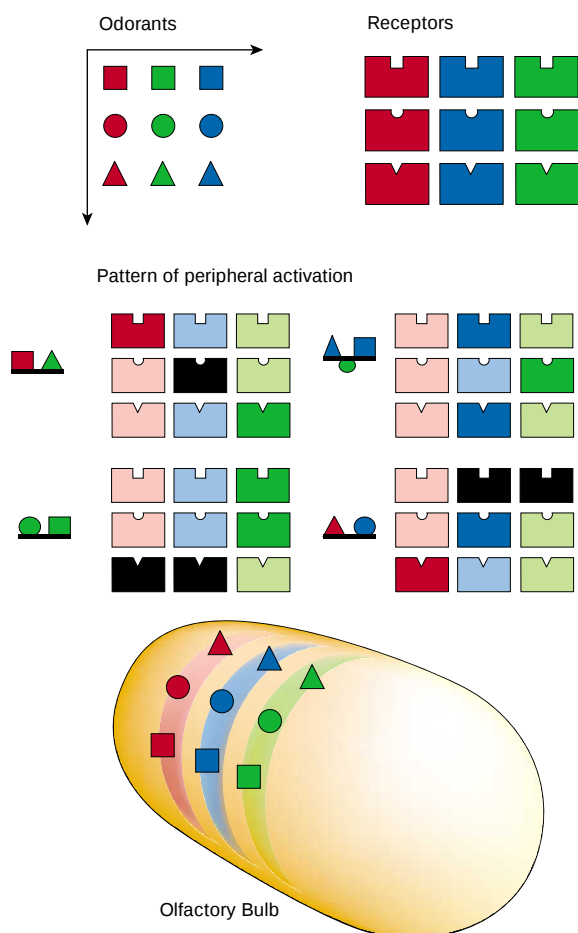
The second-messenger cascade of enzymes provides amplification and integration of odour-binding events. One membrane receptor activated by a bound odour can in turn activate tens of G proteins, each of which will activate a cyclase molecule capable of producing about a thousand molecules of cAMP per second. Three cAMP molecules are required to open a channel, but hundreds of thousands of ions can cross the membrane through a single open channel. It seems that a single odour molecule can produce a measurable electrical event in an OSN (although probably not a perceivable event in the brain) and just a few channels opening together could pass sufficient current to induce action-potential generation<sup>23,24</sup>.

Appended to this pathway is an additional, and somewhat unique, amplification mechanism in OSNs. The calcium ions entering through the CNG channel are able to activate another ion channel that is permeable to the negatively charged chloride ion<sup>25</sup>. Neuronal  $Cl^-$  channels normally mediate inhibitory responses, as  $Cl^-$  ions tend to be distributed in such a way that they will enter the cell through an open channel. But OSNs maintain an unusually high intracellular  $Cl^-$  concentration (presumably by the action of a membrane pump) such that there is a  $Cl^-$  efflux when these channels are activated. Left behind is a net positive charge on the membrane that further depolarizes the cells, thus adding to the excitatory response magnitude. This is an interesting evolutionary adaptation to the fact that the olfactory cilia reside in the mucus, outside the body proper, and where the concentrations of ions are not as well regulated as they are in normal interstitial compartments<sup>26,27</sup>. Thus the OSN maintains its own  $Cl^-$  battery, in case the  $Na^+$  gradient in the mucus is insufficient to support a threshold current, and uses it to boost the response.

Calcium ions entering through the CNG channels are also important in response adaptation through a negative feedback pathway involving the ion channel<sup>28</sup>. As intracellular calcium increases during the odour response, it acts on the channel (probably with calmodulin) to decrease its sensitivity to cAMP, thereby requiring a stronger odour stimulus to produce sufficient cAMP to open the channel<sup>29–31</sup>. This adaptation response is critical, as physiological recordings from OSNs indicate that they have very steep concentration–response relations (typically, responses from 10–90% of the maximum occur

## Box 1

## A code in the nose



Although there are some 1,000 ORs, detecting the enormous repertoire of odours requires a combinatorial strategy. Most odour molecules are recognized by more than one receptor (perhaps by dozens) and most receptors recognize several odours, probably related by chemical property. The scheme in the figure represents a current consensus model. There are numerous molecular features, two of which are represented here by colour and shape. Receptors are able to recognize different features of molecules, and a particular odour compound may also consist of a number of these 'epitopes' or 'determinants' that possess some of these features. Thus the recognition of an odorant molecule depends on which receptors are activated and to what extent, as shown by the shade of colour (black represents no colour or shape match and thus no activation). Four odour compounds are depicted with the specific array of receptors each would activate. Note that there are best receptors (for example, red square), but also other receptors that are able to recognize some feature of the molecule (for example, any square) and would participate in the discrimination of that compound. In the olfactory bulb there seem to be wide areas of sensitivity to different features (for example, functional group or molecular length). This model is based on current experimental evidence, but is likely to undergo considerable revision as more data become available.

over a stimulus concentration of about one log unit<sup>32</sup>). This means that the cells are particularly sensitive to small changes in concentration, but without adaptation to re-set the gain they would be able to respond only over a narrow dynamic range<sup>33</sup>. This is just one of several mechanisms that OSNs use for adjusting their sensitivity. Others include a recently discovered RGS (regulator of G-protein signalling) protein that apparently acts on the adenyl cyclase to decrease its activity<sup>34</sup>, and a kinase that phosphorylates activated receptors sending them into a desensitized state<sup>35,36</sup>.

Signal transduction and generation in invertebrates is far less well understood, possibly because there is not a single system at work. In lobsters, a lipid pathway involving inositol phosphates seems to be dominant<sup>37</sup>, and in *Drosophila* and the moth, there is also strong evidence for inositol-1,4,5-trisphosphate (Ins(1,4,5)P<sub>3</sub>) as a second messenger<sup>38</sup>. The difficulty in these systems is that the final target of the cascade, the analogue of the CNG channel in vertebrates, remains to be identified. In contrast to vertebrates, invertebrates have both excitatory and inhibitory responses to odours, so there are likely to be multiple transduction pathways<sup>39,40</sup>.

### Tuning curves in primary sensory cells

A classical means of describing and classifying primary sensory neurons uses the concept of a 'tuning curve' — the relationship between stimulus quality and response. For photoreceptors this would be a plot showing the range of wavelengths of light at which they are activated, and for auditory receptors it would be a similar representation of frequency. But for OSNs this is a somewhat more difficult task as

odours vary along multiple dimensions. To understand the rules of olfactory stimulus encoding it is useful to attempt to describe a 'molecular receptive range'<sup>41</sup> for OSNs.

Several approaches to this problem have been taken (Box 1). What has become clear is that most if not all receptors can be activated by multiple odours, and conversely most odours are able to activate more than one type of receptor. But this vast combinatorial strategy only underscores the importance of understanding how broadly tuned a particular OR may be. Limited physiological recordings from individual cells have produced conflicting data, probably because there are a range of tuning profiles, from specific to broad, and physiological experiments generally use necessarily limited sets of odours<sup>42</sup>. The standard means of characterizing the molecular range of a receptor is through a medicinal chemistry approach that seeks to define a pharmacophore, that is, the molecular determinants that are common to a set of agonists or antagonists for a given receptor.

Taking this approach, Araneda *et al.*<sup>43</sup> were able to provide such a characterization for at least one particular odour receptor in the rat. After screening an extensive panel of compounds they were able to determine three critical chemical features common to agonists at this receptor, and also determined that a related structural compound could serve as an antagonist, reducing the response to a known agonist. This indicates that standard pharmacology could indeed be applied profitably to the analysis of odour receptors, although large-scale screens for 1,000 ORs might be a task better suited to industrial rather than academic laboratories. It may also mean that blockers for specific malodours could be found or synthesized.



An alternative strategy would be intrinsic imaging in the olfactory bulb, which would rely on all of the OSNs expressing a particular receptor and converging onto a single glomerulus. Taking advantage of this feature, several groups<sup>44–46</sup> have recently used such optical imaging in the living olfactory bulb of rodents (see review by Dudai<sup>47</sup>). These experiments confirm that a given odour activates a set of glomeruli (that is, OSNs, and hence ORs) and that different odours activate overlapping but non-identical patterns of glomeruli (receptors). In one particularly striking case involving enantiomers (compounds of identical molecular composition that differ only in the three-dimensional arrangement of their atomic groupings) that can be discriminated behaviourally by rats, the pattern of glomerular activity induced by either stereoisomer differed by as little as a single glomerulus<sup>48</sup>. Such studies suggest that receptors that recognize similar odours tend to map in the same general area in the olfactory bulb, although at present this is merely a postulate as too few odours have been screened and only ~20% of the bulb can be visualized with this method.

### Encoding other features of the olfactory stimulus

Olfactory thresholds measured behaviourally in animals, or psychophysically in humans, are often lower than what is seen in single-cell recordings. There may be two underlying causes for this. One is the convergence at the glomerular layer described above, allowing each mitral cell to sample a large population of identically tuned primary neurons, sending even weak messages through to the brain. But there is also a cellular source for the discrepancy. Sensitivity is usually measured psychophysically as a threshold level of stimulus at which a response occurs more often than chance. At the cellular level, sensitivity is measured as the  $EC_{50}$ , or midpoint, on a dose–response curve; that is, the concentration of odour that elicits a half-maximal response. To bring these two measures together it is helpful to recognize that OSNs cannot really measure concentration, which is something of an abstraction, but rather they operate as molecular counters, tallying each interaction between an odour molecule and a receptor. A measure of the stimulus that better captures this notion would be flux, or concentration over time. With the introduction of a temporal dimension the importance of the second messenger as an integrator, as well as amplifier, becomes apparent. The second-messenger system allows the OSN to sum or integrate many individual binding events over some period of time (which has been measured at ~1 s in salamanders, and is probably shorter in mammals).

What is the value of this? For many odours the dose–response curves in single cells have relatively elevated  $EC_{50}$  values — in the range 10–100  $\mu$ M. This seems high in comparison to other GPCRs, in particular neurotransmitter receptors, but the task of the OR is different from that of a serotonin receptor. ORs are broadly tuned so as to be able to recognize a number of related but not identical molecules; this is what gives the system its tremendous range. But broadly tuned receptors cannot also have high affinities. By counting molecules and integrating over relatively long times, OSNs are able to include even low-probability binding events in generating their response. In effect the system gives up affinity for a broader receptive range, but recovers at least some of that lost sensitivity by giving up temporal resolution and using a long integration time. This seems a fair trade-off as the olfactory system is rarely called on to act quickly, as the visual or auditory systems might be.

How many odours can we detect? The literature is replete with numbers ranging from ~2,000 to more than 100,000. Theoretically it could be billions, based on the possible combinations of 1,000 receptors. In fact, the question is probably not relevant, just as it makes little sense to ask how many colours or hues we can see. Perfumers, chefs, *sommeliers* or highly trained animals are likely able to discriminate more odours than the rest of us, but this is not due to an inherent difference in equipment. Physical chemistry may be the primary limiting factor as odour chemicals must possess a certain volatility, solubility and stability to act on the nasal sensory tissue.

Another issue that has yet to be resolved concerns the effects of intensity on olfactory coding. It is often remarked that some odours change their perceived quality depending on the stimulus intensity. For example, thiols, which smell unbearably awful at high concentrations, have a sweet citrus aroma at lower concentrations. But this is far less remarkable than the fact that most odours remain constant in their quality over orders of magnitude of concentration. Amyl acetate, a pleasant fruity smelling substance, can be easily identified at concentrations from 0.1  $\mu$ M to 10 mM. By monitoring activity in the olfactory bulb it is clear that as the concentration of an odour is increased, additional glomeruli are recruited into the pattern of activity, suggesting that new receptors are being activated as concentrations increase<sup>45</sup>. Precisely how the percept remains constant as new receptors are recruited into the response is not clear. One possibility is that there may be a class of broadly tuned low-affinity receptors that are simply intensity detectors. That is, they are activated by a large number of substances, but only at higher concentrations, so that their introduction into the pattern of activity signals only a higher concentration of whatever odour the rest of the pattern was signalling.

### Pheromones and the vomeronasal system

In many mammals there is an accessory olfactory system located in a cigar-shaped organ (the vomeronasal organ or VNO) separate from the main olfactory epithelium (MOE). The VNO has been identified with the action of pheromones, molecules produced and emitted by other members of the same species. Pheromones have been implicated in mating, suckling, courtship and other behaviours and are believed to interact, through the VNO, with the endocrine system.

Two additional families of GPCRs, again unrelated to the family of ORs, have been identified in the VNO (Fig. 2), and they are differentially expressed in two segregated populations of vomeronasal sensory neurons (VSNs)<sup>49–52</sup>. Those located in the most apical portion of the epithelium express the  $G_i$  type of G protein, whereas those in the basal portion are  $G_o$  positive<sup>53</sup>. Although there is no evidence that these G proteins are involved in sensory transduction, the two families of receptors are distributed in precise coincidence with them<sup>52,54</sup>. Thus the  $G_i$ -positive neurons express receptors of the V1R family, whereas the V2R receptors are expressed in  $G_o$ -positive cells. V1R receptors number about 150 and are of the same general type of GPCR as the ORs (that is, they have short N termini). V2Rs on the other hand are similar to metabotropic glutamate receptors in that they have a long extracellular N-terminal region believed to be involved in ligand binding. There are estimated to be some 150 V2Rs in rodents, arrayed into several sub-families<sup>55,56</sup>. The organization of the vomeronasal system is somewhat different from the MOE as at least some VSNs may express more than one receptor<sup>57</sup>. VSNs project their axons to a caudal region of the olfactory bulb known as the accessory olfactory bulb (AOB; Fig. 1a), but VSN axons do not converge onto single glomeruli as in the main bulb<sup>58,59</sup>. In the AOB, sensory neurons expressing the same V1R converge on the same glomeruli, but as many as 10–30 glomeruli receive input from a given receptor (as opposed to the 1–3 in the main bulb).

With the exception of one pseudogene, none of the VRs seem to be expressed differentially in males or females<sup>50,52</sup>, indicating that sexually divergent responses to pheromones are the result of higher brain function, and that both sexes can detect all pheromones. Indeed, it may be important for a female to know that another receptive female is in the area, even though her response will be quite different from that of a nearby male.

The transduction mechanism in VSNs is not yet known, but recent physiological and biochemical evidence implicates phospholipase C and a lipid pathway possibly including  $Ins(1,4,5)P_3$ , diacyl glycerol, calcium release and perhaps the involvement of a transient receptor potential (TRP)-like calcium channel, similar to that identified in *Drosophila* phototransduction (see refs 54, 60, 61, and review this issue by Hardie and Raghu, pages 186–193). One difficulty in

obtaining these data is the lack of well characterized stimuli. Pheromones are typically a minor component of excreted fluids such as urine or sweat, and are therefore difficult to identify, purify and obtain in quantity. In at least one recent physiological study, less than 0.5% of VSNs responded to any one of a group of six putative mouse pheromones<sup>61</sup>. But with over 200 receptors expressed in the VNO, other substances besides pheromones might also be stimuli. Using calcium imaging to observe responses, Buck and colleagues measured VSN responses to 18 out of a panel of 82 common odours<sup>62</sup>. Perhaps compounds associated with the fluids containing pheromones are also sensed by this accessory olfactory system. Structurally the V2R class of receptors are candidates for binding amino acids or small peptides, as does the related mGluR. In fish, an OR similar to the V2R family was shown to bind the amino acid arginine with a high affinity<sup>14</sup>, although similar responses have not yet been shown in VSNs.

Unlike the MOE, responses of VSNs to pheromones have been observed at concentrations as low as 0.1 nM, and even at higher concentrations VSNs seem to remain highly specific for particular compounds<sup>61</sup>. This suggests that these high-affinity pathways do not use a combinatorial mechanism for sensory coding.

Are humans sensitive to pheromones? The VNO in humans is vestigial, disappearing before birth. In the human genome all putative members of the VR family are pseudogenes, with one exception. A single V1R gene was found to be intact and complementary DNA for this gene was recovered from 11 individuals of varying ethnic background. No ligand is known for this receptor nor is it known where it is expressed<sup>63</sup>. There are various behavioural studies that implicate putative pheromones in regulating endocrine-dependent behaviours such as menstruation, but the precise site of action is unknown. For an excellent recent review of the current data on this issue, see ref. 64.

## Summary and future directions

All cells, no matter what their principal function might be, need to sense and respond to important molecules in their environment. But in the olfactory system, molecular sensing is the primary occupation, and is most highly developed. During the past decade the olfactory system has emerged from relative obscurity as an idiosyncratic sensory system to its current place of interest as a model system for molecular detection. With a now firm foundation of molecular, physiological and chemical data, we can turn to the next generation of issues.

For many questions in olfaction the tools are also the puzzles. The large receptor repertoires will tell us much about GPCR structure–function relations, but how these receptors are controlled and how their expression is regulated remains mysterious. The enormous selection of ligands provides a varied chemical catalogue of agonists and antagonists, but the variety and range are bewildering and frustrate attempts at easy pharmacological-like classification. The gene families of ‘genome project species’ (for example, the nematode, fly, mouse and human) provide complete sets of OR sequences, but the challenge will be to integrate this with developmental and behavioural data. The organization of the sensory neuron inputs to the olfactory bulb is perhaps the first neural circuit to be described by molecular-genetic techniques, but the ‘meaning’ of the stimulus will not be clear until we have a better idea of what subsequent physiological processing occurs in the bulb and in the higher cortical centres.

The olfactory system is clearly ripe with opportunities for the imaginative investigator. While a sound footing has been provided by molecular, physiological, genetic, developmental and computational work over the past decade, this has all been as a prelude to addressing some of the most compelling problems in neuroscience, among them that fundamental human question: How do I smell? □

1. Hildebrand, J. G. & Shepherd, G. M. Mechanisms of olfactory discrimination: converging evidence for common principles across phyla. *Annu. Rev. Neurosci.* **20**, 595–631 (1997).
2. Buck, L. B. The molecular architecture of odor and pheromone sensing in mammals. *Cell* **100**,

- 611–618 (2000).
3. Mombaerts, P. Seven-transmembrane proteins as odorant and chemosensory receptors. *Science* **286**, 707–711 (1999).
4. Mombaerts, P. *et al.* Visualizing an olfactory sensory map. *Cell* **87**, 675–686 (1996).
5. Buck, L. & Axel, R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**, 175–187 (1991).
6. Sosinsky, A., Glusman, G. & Lancet, D. The genomic structure of human olfactory receptor genes. *Genomics* **70**, 49–61 (2000).
7. Pilpel, Y. & Lancet, D. The variable and conserved interfaces of modeled olfactory receptor proteins. *Protein Sci.* **8**, 969–977 (1999).
8. Kobilka, B. K. *et al.* Chimeric  $\alpha 2$ - $\beta 2$ -adrenergic receptors: delineation of domains involved in effector coupling and ligand binding specificity. *Science* **240**, 1310–1316 (1988).
9. Palczewski, K. *et al.* Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289**, 739–745 (2000).
10. Krautwurst, D., Yau, K. W. & Reed, R. R. Identification of ligands for olfactory receptors by functional expression of a receptor library. *Cell* **95**, 917–926 (1998).
11. Gimelbrant, A. A., Haley, S. L. & McClintock, T. S. Olfactory receptor trafficking involves conserved regulatory steps. *J. Biol. Chem.* **1**, 1 (2000).
12. Zhao, H. *et al.* Functional expression of a mammalian odorant receptor. *Science* **279**, 237–242 (1998).
13. Wetzel, C. H. *et al.* Specificity and sensitivity of a human olfactory receptor functionally expressed in human embryonic kidney 293 cells and *Xenopus laevis* oocytes. *J. Neurosci.* **19**, 7426–7433 (1999).
14. Specia, D. J. *et al.* Functional identification of a goldfish odorant receptor. *Neuron* **23**, 487–498 (1999).
15. Sengupta, P., Chou, J. H. & Bargmann, C. I. *odr-10* encodes a seven transmembrane domain olfactory receptor required for responses to the odorant diacetyl. *Cell* **84**, 899–909 (1996).
16. Clyne, P. J. *et al.* A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* **22**, 327–338 (1999).
17. Vossahl, L. B., Amrein, H., Morozov, P. S., Rzhesky, A. & Axel, R. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* **96**, 725–736 (1999).
18. Vossahl, L. B., Wong, A. M. & Axel, R. An olfactory sensory map in the fly brain. *Cell* **102**, 147–159 (2000).
19. Firestein, S., Darrow, B. & Shepherd, G. M. Activation of the sensory current in salamander olfactory receptor neurons depends on a G protein-mediated cAMP second messenger system. *Neuron* **6**, 825–835 (1991).
20. Wong, S. T. *et al.* Disruption of the type III adenylyl cyclase gene leads to peripheral and behavioral anosmia in transgenic mice. *Neuron* **27**, 487–497 (2000).
21. Belluscio, L., Gold, G. H., Nemes, A. & Axel, R. Mice deficient in  $G_{\alpha ol}$  are anosmic. *Neuron* **20**, 69–81 (1998).
22. Brunet, L. J., Gold, G. H. & Ngai, J. General anosmia caused by a targeted disruption of the mouse olfactory cyclic nucleotide-gated cation channel. *Neuron* **17**, 681–693 (1996).
23. Lynch, J. W. & Barry, P. H. Action potentials initiated by single channels opening in a small neuron (rat olfactory receptor). *Biophys. J.* **55**, 755–768 (1989).
24. Menini, A., Picco, C. & Firestein, S. Quantal-like current fluctuations induced by odorants in olfactory receptor cells. *Nature* **373**, 435–437 (1995).
25. Kleene, S. J. & Gesteland, R. C. Calcium-activated chloride conductance in frog olfactory cilia. *J. Neurosci.* **11**, 3624–3629 (1991).
26. Lowe, G. & Gold, G. H. Nonlinear amplification by calcium-dependent chloride channels in olfactory receptor cells. *Nature* **366**, 283–286 (1993).
27. Kurahashi, T. & Yau, K. W. Co-existence of cationic and chloride components in odorant-induced current of vertebrate olfactory receptor cells. *Nature* **363**, 71–74 (1993).
28. Kurahashi, T. & Menini, A. Mechanism of odorant adaptation in the olfactory receptor cell. *Nature* **385**, 725–729 (1997).
29. Chen, T. Y. & Yau, K. W. Direct modulation by  $Ca^{2+}$ -calmodulin of cyclic nucleotide-activated channel of rat olfactory receptor neurons. *Nature* **368**, 545–548 (1994).
30. Liu, M., Chen, T. Y., Ahamed, B., Li, J. & Yau, K. W. Calcium-calmodulin modulation of the olfactory cyclic nucleotide-gated cation channel. *Science* **266**, 1348–1354 (1994).
31. Kramer, R. H. & Siegelbaum, S. A. Intracellular  $Ca^{2+}$  regulates the sensitivity of cyclic nucleotide-gated channels in olfactory receptor neurons. *Neuron* **9**, 897–906 (1992).
32. Firestein, S., Picco, C. & Menini, A. The relation between stimulus and response in olfactory receptor cells of the tiger salamander. *J. Physiol.* **468**, 1–10 (1993).
33. Reiser, J. & Matthews, H. R. Adaptation-induced changes in sensitivity in frog olfactory receptor cells. *Chem. Senses* **25**, 483–486 (2000).
34. Sinnarajah, S. *et al.* RGS2 regulates signal transduction in olfactory neurons by attenuating activation of adenylyl cyclase III. *Nature* **409**, 1051–1055 (2001).
35. Dawson, T. M. *et al.* Beta-adrenergic receptor kinase-2 and beta-arrestin-2 as mediators of odorant-induced desensitization. *Science* **259**, 825–829 (1993).
36. Schleicher, S., Boekhoff, I., Arriza, J., Lefkowitz, R. J. & Breer, H. A  $\beta$ -adrenergic receptor kinase-like enzyme is involved in olfactory signal termination. *Proc. Natl. Acad. Sci. USA* **90**, 1420–1424 (1993).
37. Fadool, D. A. & Ache, B. W. Plasma membrane inositol 1,4,5-trisphosphate-activated channels mediate signal transduction in lobster olfactory receptor neurons. *Neuron* **9**, 907–918 (1992).
38. Krieger, J., Mamel, M. & Breer, H. Elements of the olfactory signaling pathways in insect antennae. *Invert. Neurosci.* **3**, 137–144 (1997).
39. Boekhoff, I., Michel, W. C., Breer, H. & Ache, B. W. Single odors differentially stimulate dual second messenger pathways in lobster olfactory receptor cells. *J. Neurosci.* **14**, 3304–3309 (1994).
40. Ache, B. W. & Zhainazarov, A. Dual second-messenger pathways in olfactory transduction. *Curr. Opin. Neurobiol.* **5**, 461–466 (1995).
41. Mori, K. & Yoshihara, Y. Molecular recognition and olfactory processing in the mammalian olfactory system. *Prog. Neurobiol.* **45**, 585–619 (1995).
42. Duchamp-Viret, P., Chaput, M. A. & Duchamp, A. Odor response properties of rat olfactory receptor neurons. *Science* **284**, 2171–2174 (1999).
43. Araneda, R. C., Kini, A. D. & Firestein, S. The molecular receptive range of an odorant receptor. *Nature Neurosci.* **3**, 1248–1255 (2000).
44. Meister, M. & Bonhoeffer, T. Tuning and topography in an odor map on the rat olfactory bulb. *J. Neurosci.* **21**, 1351–1360 (2001).
45. Rubin, B. D. & Katz, L. C. Optical imaging of odorant representations in the mammalian olfactory bulb. *Neuron* **23**, 499–511 (1999).
46. Uchida, N., Takahashi, Y. K., Tanifuji, M. & Mori, K. Odor maps in the mammalian olfactory bulb:

- domain organization and odorant structural features. *Nature Neurosci.* **3**, 1035–1043 (2000).
47. Dudai, Y. The smell of representations. *Neuron* **23**, 633–635 (1999).
48. Rubin, B. D. & Katz, L. C. Spatial coding of enantiomers in the rat olfactory bulb. *Nature Neurosci.* **4**, 355–356 (2001).
49. Dulac, C. & Axel, R. A novel family of genes encoding putative pheromone receptors in mammals. *Cell* **83**, 195–206 (1995).
50. Matsunami, H. & Buck, L. B. A multigene family encoding a diverse array of putative pheromone receptors in mammals. *Cell* **90**, 775–784 (1997).
51. Ryba, N. J. & Tirindelli, R. A new multigene family of putative pheromone receptors. *Neuron* **19**, 371–379 (1997).
52. Herrada, G. & Dulac, C. A novel family of putative pheromone receptors in mammals with a topographically organized and sexually dimorphic distribution. *Cell* **90**, 763–773 (1997).
53. Jia, C. & Halpern, M. Subclasses of vomeronasal receptor neurons: differential expression of G proteins ( $G_{\alpha_{v1}}$  and  $G_{\alpha_{v2}}$ ) and segregated projections to the accessory olfactory bulb. *Brain Res.* **719**, 117–128 (1996).
54. Krieger, J. *et al.* Selective activation of G protein subtypes in the vomeronasal organ upon stimulation with urine-derived compounds. *J. Biol. Chem.* **274**, 4655–4662 (1999).
55. Del Punta, K., Rothman, A., Rodriguez, I. & Mombaerts, P. Sequence diversity and genomic organization of vomeronasal receptor genes in the mouse. *Genome Res.* **10**, 1958–1967 (2000).
56. Pantages, E. & Dulac, C. A novel family of candidate pheromone receptors in mammals. *Neuron* **28**, 835–845 (2000).
57. Martini, S., Silvotti, L., Shirazi, A., Ryba, N. J. & Tirindelli, R. Co-expression of putative pheromone receptors in the sensory neurons of the vomeronasal organ. *J. Neurosci.* **21**, 843–848 (2001).
58. Rodriguez, I., Feinstein, P. & Mombaerts, P. Variable patterns of axonal projections of sensory neurons in the mouse vomeronasal system. *Cell* **97**, 199–208 (1999).
59. Belluscio, L., Koentges, G., Axel, R. & Dulac, C. A map of pheromone receptor activation in the mammalian brain. *Cell* **97**, 209–220 (1999).
60. Holy, T. E., Dulac, C. & Meister, M. Responses of vomeronasal neurons to natural stimuli. *Science* **289**, 1569–1572 (2000).
61. Leinders-Zufall, T. *et al.* Ultrasensitive pheromone detection by mammalian vomeronasal neurons. *Nature* **405**, 792–796 (2000).
62. Sam, M. *et al.* Odorants may arouse instinctive behaviours. *Nature* **412**, 142 (2001).
63. Rodriguez, I., Greer, C. A., Mok, M. Y. & Mombaerts, P. A putative pheromone receptor gene expressed in human olfactory mucosa. *Nature Genet.* **26**, 18–19 (2000).
64. Meredith, M. Human vomeronasal organ function: a critical review of best and worst cases. *Chem. Senses* **26**, 433–445 (2001).

#### Acknowledgements

Material for figures was provided by X. Zhang and A. Kini. Thoughtful comments were provided by members of my laboratory and D. Kelley.



# Receptors and transduction in taste

Bernd Lindemann

Department of Physiology, Universität des Saarlandes, D-66421 Homburg, Germany (e-mail: phblin@uniklinik-saarland.de)

**Taste is the sensory system devoted primarily to a quality check of food to be ingested. Although aided by smell and visual inspection, the final recognition and selection relies on chemoreceptive events in the mouth. Emotional states of acute pleasure or displeasure guide the selection and contribute much to our quality of life. Membrane proteins that serve as receptors for the transduction of taste have for a long time remained elusive. But screening the mass of genome sequence data that have recently become available has provided a new means to identify key receptors for bitter and sweet taste. Molecular biology has also identified receptors for salty, sour and umami taste.**

**T**here is no life form known between bacteria and mammals that would neglect to check its intake of matter by chemoreceptive scrutiny. A human baby, only a few days old, already can distinguish sweet and bitter and express pleasure for sweet taste but displeasure for bitter taste<sup>1</sup>. Inorganic ions, sugars and polysaccharides, amino acids and peptides, toxins and 'xenobiotics' are all subject to nutritional chemoreception followed by adaptive behaviour. But details differ widely, depending on the complexity of the organism and the ecological niche that it occupies. Even in closely related species, distinct differences in sensory performance may be noted, which seem to match the nutritional 'needs' of a species. To understand how such a match arose, that is, how receptor specificity changed with the availability of food ingredients, is perhaps the most fascinating of the future tasks of taste research.

## Taste buds and taste cells

Already in worms, like the model nematode *Caenorhabditis elegans*, a distinction can be made between olfaction and taste<sup>2</sup>. These two chemoreceptive senses are more clearly separate in arthropods and they are quite distinct in vertebrates. In the fruitfly *Drosophila melanogaster*, for example, taste sensations are mediated by nerve cells of characteristic topology. Their sensory dendrites are contained in 'hairs' found on the body surface. Other taste neurons, found on the labellum and clustered around the pharynx, express a family of G-protein-coupled receptors (GPCRs) named GR<sup>3</sup>. This family, however, contains candidate receptors for both taste and olfaction, as its genes are expressed in both gustatory and olfactory primary neurons<sup>4</sup>. In contrast, the taste receptor cells of vertebrates are not neurons, but originate from the epithelial covering of the body<sup>5</sup>.

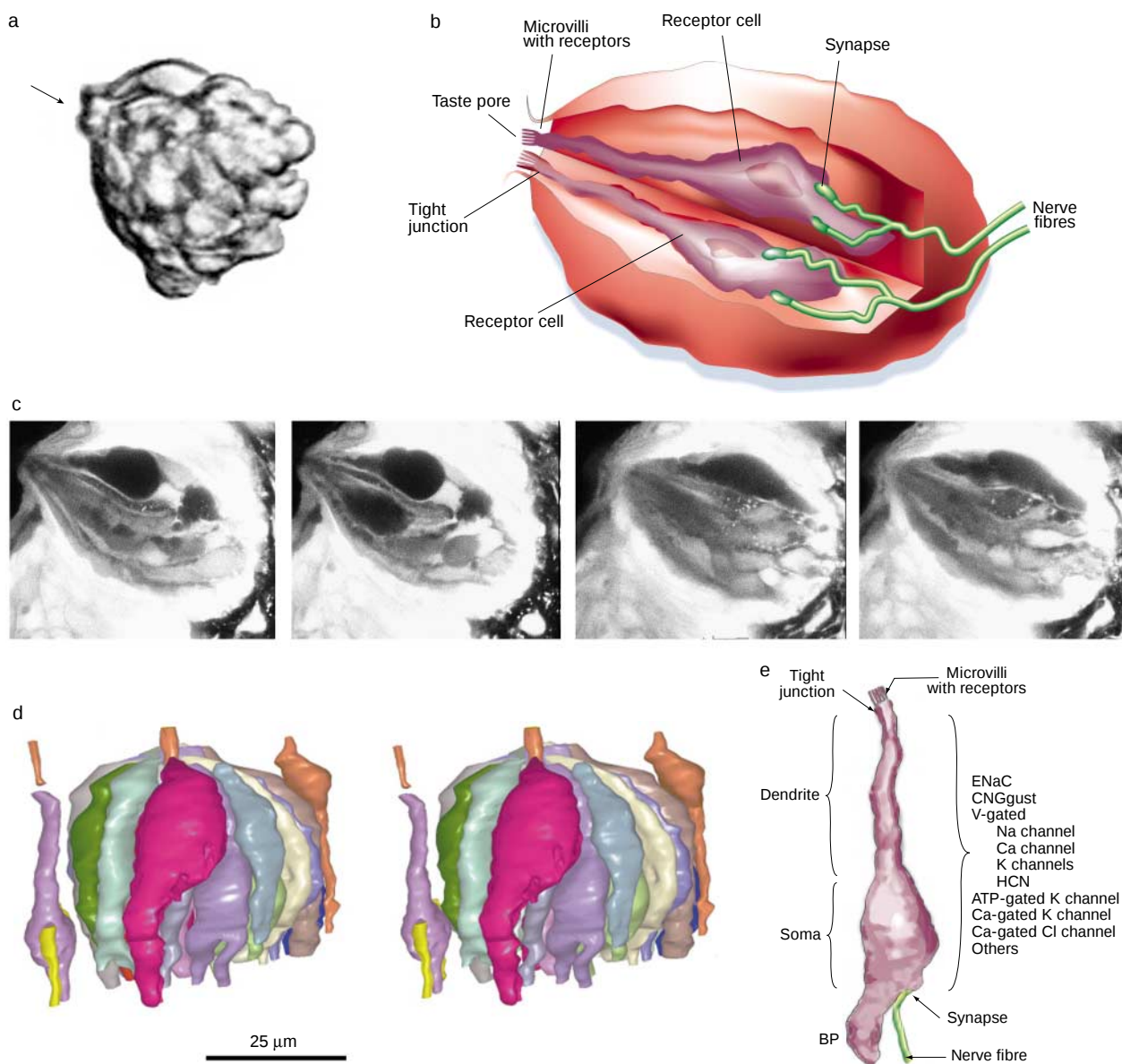
Vertebrate taste cells are small bipolar cells (Fig. 1). To connect to the oral space, they send a thin dendritic process to the epithelial surface. The cells occur either singly or densely packed in taste buds, where up to 100 form a functional unit (Fig. 1a, d). Although taste buds also occur abundantly on the body surface and barbels of some fish, all vertebrates have taste buds in the oral epithelium, typically on tongue, palate and pharynx. On the tongue, the taste buds are mounted in special folds and protrusions called papillae. The marker molecule gustducin, a taste-specific G protein<sup>6</sup>, shows additional 'taste cells' in the nasal mucosa<sup>7</sup> and in the stomach<sup>8</sup>.

A century ago it was determined that each chemoreceptive area of the human tongue responds to each of the qualities of sweet, sour, salty and bitter taste. Only minor differences in subjective thresholds were noted across areas<sup>9,10</sup>. May the time come soon when textbook authors wake up to this fact! In rodents, the differences in thresholds or sensitivities seem to be somewhat larger<sup>11</sup>. Here, morphological and functional differentiation between buds from the anterior and posterior tongue can be noted more easily, even though individual taste buds of all areas contain cells responding to several qualities.

## Lifespan, connectivity and coding

Mammalian taste-responsive cells age fast, their lifespan being about 10 days. This implies that any one nerve terminal in a taste bud frequently has to detach from an ageing cell, find a developing taste cell and form new synapses on its surface. The new cell has presumably to be of appropriate specificity such that a nerve fibre previously attached to a Na<sup>+</sup>-responsive cell, for instance, will again attach to a cell of this kind. As taste cells of different specificity occur together in one bud, one expects surface markers to guide a nerve fibre to the right cellular target. In fact, the receptor molecules themselves may possibly serve as markers, as they occur not only on the microvilli, but also on basolateral membrane areas of the receptor cells.

The basic rules for the developmental interactions that occur between taste cells and nerve fibres begin to be deduced using the methods of molecular biology<sup>12,13</sup>, and provide a foundation for a further analysis of the connectivity and the 'sensory code' by more function-oriented approaches such as electrophysiology and fluorescence imaging. Meanwhile, recordings from the sensory nerve fibres and from the soma of the neurons have consistently revealed the unexpected and striking feature that some nerve fibres are specialists, but many are generalists, carrying responses to more than one taste quality<sup>14</sup>. A simple 'labelled line' design, where each fibre responds to just one of the qualities, to bitter only or to sour only, is not evident, as many fibres are broadly tuned with respect to taste ligands. These generalist fibres carry responses to salty and sour, to glutamate and sucrose, and so on. Similarly, many taste receptor cells, too, are generalists, as responses to taste qualities are randomly and independently distributed, varying in intensity across cells<sup>15</sup>. Given such distributed responses, a part of the information about individual tastants must be buried in the quorum of the receptor cells and the 'across-fibre pattern' of the sensory nerve<sup>16</sup>. Retrieval



**Figure 1** Morphology of taste buds (rat). **a**, Viable bud isolated from the vallate papilla. Taste pore at the upper left (arrow). Length of bud is 30  $\mu\text{m}$ . **b**, Cut-open view of a bud (cartoon). Highlighted are two receptor cells with apical microvilli and basolateral synapses. **c**, Images of a viable bud from the vallate papilla, taken with a 2-photon microscope. The four optical planes depict multiple bipolar cells in different states of loading with a fluorescent dye, and nerve fibres. **d**, Three-dimensional reconstruction, from microscopic serial sections, of a bud from the foliate papilla, the taste pore facing upwards. On the left, a solitary bipolar cell with innervating nerve fibre is also visible. Scale bar, 25  $\mu\text{m}$ . (Image courtesy of V. I. Popov, Institute of Cell Biophysics, RAS, Pushchino, Russia.). **e**, Bipolar receptor cell with sensory nerve fibre attached. Some morphological details and the location of the main types of identified ion channels on the lateral membrane are indicated. BP, basal cell process.

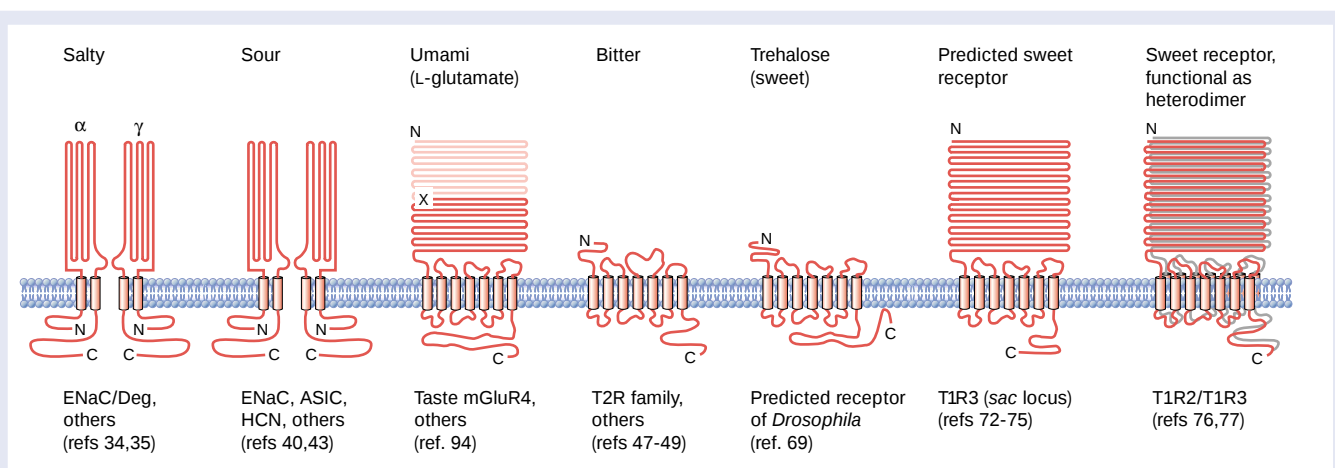
of this information, for instance by pattern discrimination, will be one of the main tasks of central taste processing.

### Receptors and transduction

The bipolar taste cells have two specializations of obvious functional significance: microvilli in contact with the oral cavity and synapses with sensory nerve fibres. Taste receptor proteins are mounted on the microvilli, acting as molecular antennas listening into the chemical environment. On binding taste molecules, the receptors trigger transduction cascades that activate synapses and thus cause excitation of the nerve fibres. These carry the signal to the brain stem, where central taste processing begins, ultimately eliciting adaptive responses.

The first molecular encounter with tastants is made therefore by those membrane proteins — the ‘receptors’ — in the apical surface of

taste receptor cells (Fig. 1b, e). They provide the molecular specificity of the taste response. A plethora of proteins, including ion channels, ligand-gated channels, enzymes and GPCRs, serve as receptors for sensory qualities such as salty, sour, sweet, umami and bitter taste (Fig. 2), and trigger the downstream transduction events within taste cells. Included among these events is the firing of action potentials, which taste cells, like neurons, are able to generate by means of voltage-gated  $\text{Na}^+$ ,  $\text{K}^+$  and  $\text{Ca}^{2+}$  channels (Fig. 1e)<sup>17–19</sup>. A local increase in  $\text{Ca}^{2+}$  concentration is needed for synaptic activation (and hence nerve excitation), and transient rises in the cytosolic  $\text{Ca}^{2+}$  concentration were observed by fluorescence imaging (Fig. 1c) in taste cells responding to bitter and sweet agents<sup>20–22</sup>, while amino acids triggered either increases or decreases of the  $\text{Ca}^{2+}$  signal<sup>23,24</sup>.



**Figure 2** Taste receptors of known primary structure, discovered 1998–2001.

A number of transmitters have been found within taste buds, but those released by taste cell synapses have been difficult to identify. Noradrenaline and acetylcholine seem to be secreted by nerve fibres and modulate the responses of taste cells<sup>25,26</sup>. Serotonin is thought to act as a paracrine agent between taste cells. Secreted by one cell and modulating the taste response of a neighbouring cell, this agent mediates local signal processing within a taste bud<sup>27,28</sup>. Glutamate is a strong candidate for a mainstream afferent transmitter secreted by taste cell synapses<sup>29,30</sup>.

### Salt taste

Two taste qualities detect ions in the oral space: salt and sour taste. Salt taste guides the incorporation of NaCl and other required minerals, thus serving an essential function in ion and water homeostasis. It shows variations across animal species, depending on the ion content of the characteristic diet<sup>19</sup>.

Although salt taste is elicited by many ionic species, the component due to the presence of Na<sup>+</sup> ions may be the most relevant for mammals and is certainly the best studied<sup>31</sup>. The Na<sup>+</sup> taste, in turn, is composed of a Na<sup>+</sup>-specific and a nonspecific mechanism. It was long suspected that a sodium channel sensitive to the channel-blocker amiloride serves as a salt receptor<sup>32</sup>. This conjecture holds especially for rodents, where the Na<sup>+</sup>-specific salt taste is indeed mediated by a highly Na<sup>+</sup>-selective channel known as ENaC, the amiloride-sensitive epithelial sodium channel<sup>33</sup>. ENaC is a hetero-oligomeric complex comprising three homologous subunits (Fig. 2), which acts as a salt-taste receptor by providing a specific pathway for sodium current into taste cells, provided that Na<sup>+</sup> ions are present in the oral space in sufficient concentration<sup>34,35</sup>. The current triggers action potentials at the basolateral membrane of the taste cell<sup>36</sup>, followed by synaptic events.

Of the three essential subunits of ENaC, at least one is under inductive control by a steroid hormone, aldosterone<sup>35</sup>. Thus the sensitivity of sodium taste is increased in animals in sodium-need through induction of more ENaC channels. A systemic Na<sup>+</sup> deficiency, which leads to salt craving, occurs regularly in herbivores, but can be observed also in rodents and humans. The induction of ENaC subunits in vallate taste buds by circulating aldosterone provides an instructive example for adaptive tuning of taste acuity in a state of nutritional deficiency.

In humans, the amiloride sensitivity of salt taste is less pronounced<sup>37</sup>, suggesting the involvement of another, as yet unspecified channel. It is ironic that so little is known especially about the molecular foundation of human salt taste.

### Sour taste

Sour taste is acceptable or interesting when mild, thereby aiding the recognition of complex food, but it becomes increasingly unpleasant

when strong<sup>1</sup>. It serves to detect unripe fruits and spoiled food, and to avoid tissue damage by acids and problems of systemic acid–base regulation.

The receptors proposed for sour taste in mammals can be ordered in two groups. The first comprises channels that conduct an inward proton current if protons are available in the oral space. ENaC has this property<sup>38</sup>. The second group comprises H<sup>+</sup>-gated channels, including the apical K<sup>+</sup> channel of the mudpuppy *Necturus*<sup>39</sup>, MDEG1 of the ENaC/Deg family<sup>40</sup>, an unspecific H<sup>+</sup>-gated cation channel<sup>41,42</sup>, and HCN, the hyperpolarization-activated, cyclic nucleotide-gated cation channel<sup>43</sup>. The large variety of mechanisms found for sour taste highlights the complexity of taste transduction.

To some extent, the intracellular pH of taste cells follows extracellular changes in pH, especially when the extracellular changes extend over the basolateral membrane. This probably occurs because the tight junction, which closes the extracellular space of a taste bud towards the oral space (Fig. 1b), is permeable to H<sup>+</sup> ions<sup>44</sup>. H<sup>+</sup> ions may therefore invade the taste bud and initiate intracellular ‘pH tracking’, which is thought to contribute to sour transduction<sup>45</sup>.

### Bitter taste

Bitter taste is unpleasant though bearable when weak, but repulsive when strong. Many organic molecules, originating from plants and interfering with the internal signalling system of animals and humans, are bitter, including caffeine, nicotine and strychnine, and the same is true for many drugs produced by industry<sup>46</sup>. Bitter taste effectively warns us not to ingest potentially harmful compounds. One of the exciting challenges in taste research is to understand how bitter receptors were shaped by evolution to serve this task.

By scanning genomic databases near a promising bitter locus on mouse chromosome 6, a group of new GPCRs was discovered, the T2R family<sup>47,48</sup>. The receptors had short amino-terminal domains (NTDs) (Fig. 2) and were expressed specifically in a subset of taste cells. The large size of this family, with 40–80 members originally specified<sup>47</sup>, came as a surprise. Three of the receptors could be expressed in a cell line, where they responded to bitter tastants<sup>49</sup>. In humans, the T2R family comprises at least 24 genes coding for GPCRs, distributed over three chromosomes (B. Bufer and W. Meyerhof, personal communication).

It is likely but not yet certain that all 24 GPCRs of the T2R family respond to bitter agents. How are these GPCRs distributed across taste cells? Adler *et al.* concluded from *in situ* hybridization data<sup>47</sup> that ‘a single taste receptor cell expresses a large repertoire of T2Rs, suggesting that each cell may be capable of recognizing multiple tastants’. Caicedo and Roper have probed this question differently, by using functional imaging experiments. They found that most bitter-responsive taste cells, which presumably expressed T2Rs, were



activated by only one out of five compounds tested. Thus, the tuning of receptor cells with respect to bitter compounds seems to be more focused than anticipated, and different bitter stimuli may activate, through the GPCRs expressed, different subpopulations of bitter-sensitive taste cells<sup>22</sup>. In agreement with this finding, single taste nerve fibres carry signals that discriminate between bitter compounds<sup>50</sup>.

It is not only GPCRs that act as bitter receptors. Some bitter peptides with amphiphilic properties interact directly with G proteins, probably by virtue of a structural similarity to the G-protein-binding site of the receptor<sup>46</sup>. Quinine, also an amphiphilic compound, permeates the cell membrane and activates G proteins, bypassing the receptor<sup>51</sup>. In *Necturus*, quinine blocks apical K<sup>+</sup> channels<sup>52</sup>, whereas in the bullfrog, it activates a cation conductance in taste cells<sup>53</sup>. Denatonium blocks voltage-gated delayed-rectifier K<sup>+</sup> channels<sup>54</sup>.

Caffeine and other methyl-xanthines also act without activating a GPCR (Fig. 3). After permeating the cell membrane they block an intracellular phosphodiesterase and cause activation of a soluble guanylate cyclase<sup>55</sup>. The latter effect may be under control of nitric oxide, as nitric oxide synthase was found in taste cells<sup>56</sup>. As the result of these complex events, a transient increase in guanine 3',5'-cyclic monophosphate (cGMP) was measured with stopped-flow methods<sup>55</sup>.

### Bitter-taste transduction

Members of the T2R family were found co-expressed with the  $\alpha$ -subunit of the G protein gustducin<sup>47</sup>, a taste-specific signalling protein long known to have a prominent role in bitter taste<sup>6</sup>. Knockout mice lacking  $\alpha$ -gustducin show decreased sensitivity for bitter agents such as denatonium and quinine, and also for sweet agents such as saccharin and sucrose<sup>57</sup>.  $\alpha$ -Gustducin activates a taste-specific phosphodiesterase<sup>58</sup>, lowering the cellular concentration of cyclic nucleotides<sup>59</sup>.

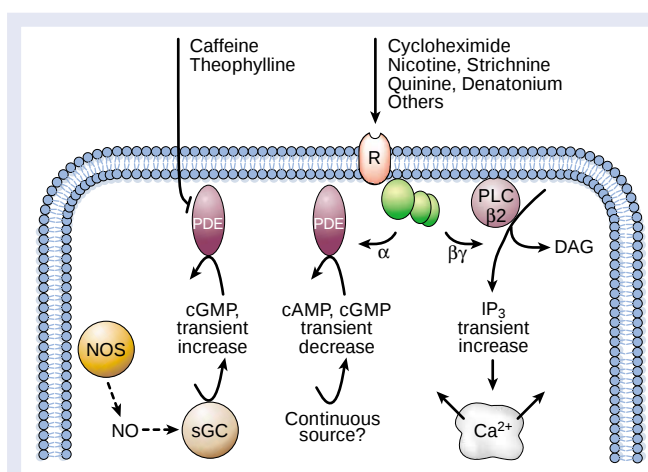
Another transduction cascade is also activated simultaneously by the GPCR-mediated bitter signal (Fig. 3). Of the  $\beta$ - and  $\gamma$ -subunits of heterotrimeric gustducin<sup>60</sup>, G $\gamma$ 13 and G $\beta$ 3 are able to activate phospholipase C $\beta$ 2 (PLC $\beta$ 2; refs 60–62), leading to the generation of inositol-1,4,5-trisphosphate (Ins(1,4,5)P<sub>3</sub>)<sup>63,64</sup>. This messenger activates Ins(1,4,5)P<sub>3</sub> receptors of intracellular Ca<sup>2+</sup> stores, of which the type III receptor, which may be modulated by calcium ions and by cAMP-dependent kinases, is the dominant form<sup>65</sup>. The receptor is co-expressed with PLC $\beta$ 2 (ref. 66). Concomitantly, Ca<sup>2+</sup> ions are released into the cytosol<sup>120,21,67</sup>. The subsequent elements of transduction, that is, the channels responsible for a change in membrane potential, remain to be identified.

Thus, the GPCR-mediated bitter signal triggers a transient decrease in cAMP and cGMP, accompanied by a transient increase in Ins(1,4,5)P<sub>3</sub> (refs 59, 64). Whether the two pathways, which are activated simultaneously, are connected, with the decrease in cyclic nucleotides enabling the increase in Ins(1,4,5)P<sub>3</sub>, has not yet been determined. Furthermore, it is unclear why dual signalling should be required for bitter taste, and whether or not it is more than parallel amplification. A similar design is not found in the receptor cells of other senses, and the other major chemoreceptive organ — the nose — seems to do well without it (see review in this issue by Firestein, pages 211–218).

### Sweet taste

Sweet taste responds to soluble carbohydrates present in sufficient concentrations in the oral cavity, guiding high-calorie intake. Yet a wide diversity of non-carbohydrate molecules is also sweet. Sweet taste has a strong hedonic (pleasant) effect<sup>1</sup>.

Considerable efforts have been made by chemists and researchers in food-producing companies to deduce from hundreds of sweet-tasting compounds common structural features that were expected to capture characteristics of 'the' sweet molecule and, therefore, 'the' sweet receptor. The binding-site models obtained seemed realistic in that they could be used to design new high-potency sweeteners<sup>68</sup>. The time has come, however, to define sweet receptors more directly using the tools of genetics and molecular biology.



**Figure 3** Transduction of bitter taste as elicited by a variety of ligands. Rs, multiple GPCRs of the T2R family, coupled to the G protein gustducin<sup>47–49</sup>,  $\alpha$ ,  $\alpha$ -subunit of gustducin<sup>61</sup>;  $\beta\gamma$ , G-protein subunits  $\beta$ 3 and  $\gamma$ 13 (refs 60–62); PLC $\beta$ 2, phospholipase C subtype<sup>61</sup>; Ins(1,4,5)P<sub>3</sub>, inositol-1,4,5-trisphosphate<sup>59</sup>; PDE, taste-specific phosphodiesterase<sup>58</sup>; cAMP, cyclic adenosine monophosphate<sup>59</sup>; cGMP, cyclic guanosine monophosphate<sup>59</sup>; sGC, soluble guanylate cyclase<sup>55</sup>; NO, nitric oxide<sup>55</sup>; NOS, NO synthase<sup>56</sup>. For second-messenger kinetics, see refs 55, 59, 63, 64.

A first success was achieved by cloning a candidate trehalose receptor of *Drosophila*<sup>69</sup>. It is a GPCR that has only a short NTD (Fig. 2). In the mouse genome, sweet-receptor genes are most likely located on chromosome 4, where two sweet taste-related locations are found, the *Dpa* locus and the *Sac* locus<sup>70,71</sup>. Mutations in the *Dpa* locus resulted in a partial loss of taste acuity for the sweet amino acid D-phenylalanine, whereas mutations in the *Sac* locus caused a partial loss of taste acuity for sucrose, saccharin and other sweeteners. Searching the new genomic databases near these two loci may uncover genes for sweet receptors.

When applied recently to the *Sac* locus, this strategy proved remarkably successful. The receptor T1R3, found simultaneously by four laboratories<sup>72–75</sup>, has a large NTD (Fig. 2), just like the orphan receptors T1R1 and T1R2 described previously<sup>76</sup>. T1R3 was found in buds of the anterior, lateral and posterior tongue. It is expressed in many taste cells which also express the orphan T1R2, suggesting that the two receptors serve a common function. They might, for instance, form heterodimers, as is known from some other GPCRs with large NTDs<sup>72,73</sup>.

As the gene *Tas1r3* is the only GPCR-coding gene at the *Sac* locus, its product T1R3 is a strong candidate for a sweet receptor. This conjecture is supported by additional observations. First, strains of mice that differ in sweet-taste ability also differ by several point mutations in *Tas1r3*. These changes do not affect expression, but may result in a decline of function. Of course, owing to extended polymorphism, the strains also differ by many more mutations outside of *Tas1r3*. Second, inbreeding of taster and non-taster strains revealed a correlation between alleles received and sweet-taste acuity<sup>73</sup>. In the future, a more detailed proof might show that transgenic non-taster mice become tasters when *Tas1r3* is switched on. The ink was not yet dry on the candidate sweet receptor T1R3, when Charles Zuker and co-workers managed to express this GPCR in oocytes. They found that the receptor does not respond to sweeteners on its own. But responses were obtained after co-expressing T1R3 together with T1R2, showing that the first functional sweet receptor found in mammals is a dimer<sup>77</sup>.

### Sweet-taste transduction

How is T1R3 coupled to the transduction machinery? This GPCR is expressed by about 20% of the taste cells, some of which also express  $\alpha$ -gustducin. The precise extent of co-expression is debatable, however, as one group found that less than 20% of the T1R3 cells have a

measurable gustducin signal<sup>73</sup>, whereas another specified a much larger percentage<sup>72</sup>. The co-expression is compatible with a role of  $\alpha$ -gustducin in sweet taste, as data obtained from knockout mice had suggested previously<sup>57</sup>. If the fraction of cells that co-express T1R3 with  $\alpha$ -gustducin is as low as 20%, then the co-expressing cells may be generalists, responding to both sweet and bitter signals. In this case,  $\alpha$ -gustducin would be a hallmark of bitter transduction, but would not be present in most sweet-responsive cells, which ignore the bitter signal. Many of the taste cells co-expressing T1R3 and  $\alpha$ -gustducin also express G $\beta$ 3, G $\gamma$ 13 and PLC $\beta$ 2 (ref. 72), all transduction elements of bitter taste (Fig. 3).

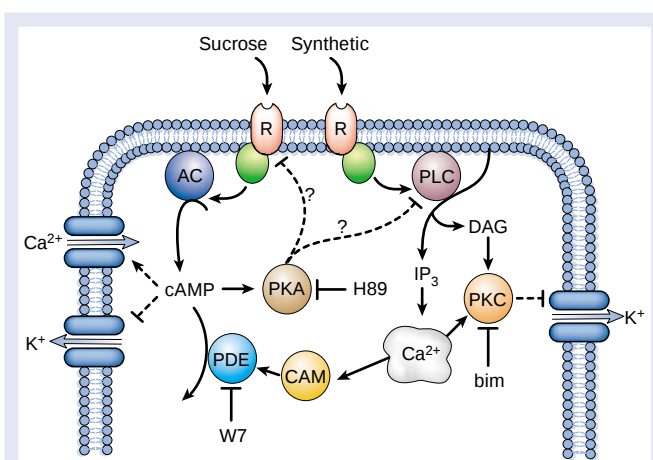
Like bitter-responsive cells, sweet-responsive cells use both cyclic nucleotides and Ins(1,4,5)P<sub>3</sub> as second messengers (Fig. 4). Ca<sup>2+</sup>-imaging experiments with isolated taste buds of rat vallate origin showed that stimulation with sugars or with forskolin caused Ca<sup>2+</sup> uptake from the extracellular space, whereas non-sugar sweeteners caused Ca<sup>2+</sup> release from intracellular stores<sup>21</sup>. This, together with other data, suggested that sugars activate a cyclic nucleotide cascade, leading to an increase of cAMP, membrane depolarization and Ca<sup>2+</sup> uptake, whereas non-sugar sweeteners activate the Ins(1,4,5)P<sub>3</sub> cascade in the same cell<sup>21</sup>. This notion was compatible with results from several laboratories, some of which involved inhibition of a K<sup>+</sup> conductance as the depolarizing step<sup>18,78–87</sup>. Inhibition of the K<sup>+</sup> conductance was thought originally to occur through protein kinase A (PKA)<sup>78</sup>, but a cyclic nucleotide-gated channel, the CNG<sub>gust</sub>, that was found in taste cells<sup>88</sup> might contribute to depolarization and Ca<sup>2+</sup> inflow when cAMP increases (Fig. 4). A co-localization of the channel with T1R3 or  $\alpha$ -gustducin has not yet been reported.

More recent evidence indicates modifications regarding the role of cAMP in sweet taste (Fig. 4). An inhibitor of PKA was found not to inhibit the sugar-sweet response in the hamster anterior tongue, but to enhance it<sup>89</sup>. This indicates that PKA is not involved directly in the response to sugars, but may be involved in adaptation. The response to artificial sweeteners was also enhanced. In contrast, inhibition of protein kinase C did not affect responses to sucrose, but inhibited responses to artificial sweeteners. This showed again that the transduction of the two kinds of sweeteners differs. Inhibition of the Ca<sup>2+</sup>/calmodulin-dependent CAMP-phosphodiesterase enhanced the responses to sucrose but not to synthetic sweeteners, indicating that the calcium ions released during stimulation with synthetic sweeteners may depress a simultaneous response to sucrose by activation of this enzyme<sup>89</sup>.

Furthermore, it was found with improved recording conditions that in the vallate papilla of the rat the second messenger cGMP rose transiently, with a peak-time of 150 ms, in response to sucrose. This signal was observed only as long as cAMP remained low, suggesting that cGMP is involved in the initial stage of sugar-taste transduction and may be more significant than cAMP at this stage<sup>90</sup>.

In conclusion, sweet-taste transduction is complex and our knowledge about it is far from complete. The data presently available suggest that sweet stimuli activate taste cells through at least two transduction pathways (Fig. 4), of which one involves an increase in cyclic nucleotides (cGMP or cAMP), the other an increase in Ins(1,4,5)P<sub>3</sub>. Membrane depolarization by inhibition of a K<sup>+</sup> conductance may be a common feature of the two pathways. An increase in the cytosolic Ca<sup>2+</sup> concentration occurs in both of the pathways, even though the source of the Ca<sup>2+</sup> ions is different. There seems to be variability in utilization of the pathways across the anterior and posterior regions of the tongue and across sweeteners and animal species. Now that candidates for sweet receptors are known, new biochemical experiments may soon identify the transduction elements downstream of the receptors.

Like other taste qualities, sweet taste is modified by circulating hormones. Recently, the effect of leptin on sweet-responding taste cells has generated much interest. Leptin, a protein hormone encoded by the *ob* gene, is secreted mainly by adipocytes and regulates body mass. The full-length leptin receptor Ob-Rb, a principal mediator of



**Figure 4** Molecules involved in the transduction of sweet taste. Two separate sweet receptors are shown, but the possibility that one receptor activates both of the transduction pathways<sup>100</sup> is not excluded at this stage. R, candidate receptor(s)<sup>72–75</sup>; AC, adenylyl cyclase<sup>81,82,87</sup>; cAMP, cyclic adenosine monophosphate<sup>21</sup>; PDE, phosphodiesterase, inhibitor W7 (ref. 89); CAM, calmodulin<sup>89</sup>; PKA, protein kinase A, inhibitor H89 (ref. 89); PLC, phospholipase C<sup>89</sup>; DAG, diacylglycerol; Ins(1,4,5)P<sub>3</sub>, inositol 1,4,5-trisphosphate<sup>21</sup>; PKC, protein kinase C, inhibitor bim (bisindolylmaleimide)<sup>89</sup>. For crosstalk between pathways and effects of inhibitors (H89, bim and W7), see ref. 89.

the leptin signal, is expressed in various tissues including pancreatic  $\beta$ -cells and a subset of taste cells<sup>91,92</sup>. Leptin suppresses insulin secretion by the activation of ATP-sensitive K<sup>+</sup> channels. Its inhibitory effect on taste receptor cells also involves the activation of a K<sup>+</sup> conductance and membrane hyperpolarization<sup>91</sup>. Thereby the hormone partially blunts nerve signals indicating sweet taste, which, presumably, makes food less attractive. During starvation the production of leptin is decreased. The resulting disinhibition in the target tissues diminishes energy expenditure and leads to the motivational state of hunger. At the same time, disinhibition of sweet-responsive taste cells enhances sensitivity to sweet taste, making sweet food more attractive. Thus the effect of leptin on the taste system supports the general role of this hormone in regulating nutrition, body weight and energy balance<sup>92</sup>.

## Umami taste

The biological significance of this basic taste, discovered about 100 years ago, is high, comparable perhaps to that of sweet taste. 'Umami', a term derived from the Japanese *umai* (delicious), designates a pleasant taste sensation which is qualitatively different from sweet, salty, sour and bitter<sup>93</sup>. Umami is a dominant taste of food containing L-glutamate, like chicken-broth, meat extracts and ageing cheese. The rather common amino acid L-glutamate thus guides the intake of peptides and proteins, from which it is released by proteolysis (curing and decay). Characteristic taste-enhancing effects arise from the presence of purine 5'-ribonucleotides such as IMP and GMP, which are also present in decaying biological tissues.

A taste receptor for L-glutamate might possibly be related to one of the glutamate receptors well known from neuronal synapses. Starting with this hypothesis, it was discovered that a subset of taste cells contains a truncated form of brain mGluR4, a metabotropic GPCR abundant in the central nervous system. Although brain mGluR4 has a rather large NTD, the taste variant has a truncated NTD (Fig. 2) and this seems to adapt the receptor to the high glutamate concentrations occurring in food<sup>94</sup>. Synergism with ribonucleotides, a highlight of umami taste, was established<sup>95</sup>. Once again, the transduction is complex — one variant involves the sustained closure of an unspecific cation conductance, presumably causing hyperpolarization, even though transient inward currents, which would cause

depolarization, were also observed<sup>196,97</sup>. In addition to the taste mGluR4, other glutamate and amino-acid receptors were also found to function in taste cells<sup>23,24,98</sup>.

### Perspectives for taste research

Much effort is now being made to identify the receptors and other key molecules of transduction within taste receptor cells. The sequencing of the receptor genetic code in particular opens the way to study the corresponding proteins and map their structure with good spatial resolution. In this way, we should achieve a better understanding of how the protein machinery within taste GPCRs actually works.

The practical consequences of these efforts are considerable. Based on binding-site structure, advanced techniques of drug design are expected to allow the construction of taste ligands that activate or inhibit a receptor protein, thereby enhancing or inhibiting a specific taste. Thus it might become possible to expand the already huge commercial market for artificial sweeteners into other taste qualities. This will be beneficial in many ways. For example, aged people often have a general decline of taste function<sup>99</sup> and need taste enhancement to once again enjoy their food. And an organic enhancer of sodium taste would be a great help for patients on a low-sodium diet.

Other scientific challenges still wait to be tackled by taste research. In the evolution of animal species, adaptive changes of taste receptors have occurred, which supported or generated new food preferences, probably driven by changes in food availability. We do not yet understand these long-term changes. The future large-scale sequencing of animal genomes may enable us to shed light on this interesting aspect of evolution.

Finally, the perspective is likely to widen in the future, and taste-driven processing in the brain, today studied by few, will be a central theme of the field. Challenging topics such as the sensory code, short-term memory of taste (often exploited as conditioned taste acceptance and aversion), hormonal feedback systems serving nutritional needs, and the generation of motivational states which guide feeding behaviour will increasingly come into focus. A deeper understanding of our conscious and unconscious decisions in food selection and food intake may then have applications in diverse fields, from food processing to clinical medicine. □

- Ganchrow, J. R., Steiner, J. E. & Daher, M. Neonatal facial expressions in response to different qualities and intensities of gustatory stimuli. *Infant Behav. Dev.* **6**, 189–200 (1983).
- Pierce-Shimomura, J. T., Faumont, S., Gaston, M. R., Pearson, B. J. & Lockery, S. R. The homeobox gene *lim-6* is required for distinct chemosensory representations in *C. elegans*. *Nature* **410**, 694–698 (2001).
- Clyne, P. J., Warr, C. G. & Carlson, J. R. Candidate taste receptors in *Drosophila*. *Science* **287**, 1830–1834 (2000).
- Scott, K. *et al.* A chemosensory gene family encoding candidate gustatory and olfactory receptors in *Drosophila*. *Cell* **104**, 661–673 (2001).
- Stone, L. M., Finger, T. E., Tam, P. P. & Tan, S. S. Taste receptor cells arise from local epithelium, not neurogenic ectoderm. *Proc. Natl Acad. Sci. USA* **92**, 1916–1920 (1995).
- McLaughlin, S. K., McKinnon, P. J. & Margolske, R. F. Gustducin is a taste-cell-specific G protein closely related to the transducins. *Nature* **357**, 563–569 (1992).
- Zancanaro, C., Caretta, C. M., Merigo, F., Cavagioni, A. & Osculati, F.  $\alpha$ -Gustducin expression in the vomeronasal organ of the mouse. *Eur. J. Neurosci.* **11**, 4473–4475 (1999).
- Höfer, D., Püschel, B. & Drenckhahn, P. H. Taste receptor-like cells in the rat gut identified by expression of  $\alpha$ -gustducin. *Proc. Natl Acad. Sci. USA* **93**, 6631–6634 (1996).
- Hänig, D. P. Zur Psychophysik des Geschmackssinnes. *Phil. Stud.* **17**, 576–623 (1901).
- Lindemann, B. Receptor seeks ligand: on the way to cloning the molecular receptors for sweet and bitter taste. *Nature Med.* **5**, 381–382 (1999).
- Smith, D. V. & Margolske, R. F. Making sense of taste. *Sci. Am.* **284**, 26–33 (2001).
- Mistretta, C. M., Goosens, K. A., Farinas, I. & Reichardt, L. F. Alterations in size, number, and morphology of gustatory papillae and taste buds in BDNF null mutant mice demonstrate neural dependence of developing taste organs. *J. Comp. Neurol.* **409**, 13–24 (1999).
- Krimm, R. F., Miller, K. K., Kitzman, P. H., Davis, B. M. & Albers, K. M. Epithelial overexpression of BDNF or NT4 disrupts targeting of taste neurons that innervate the anterior tongue. *Dev. Biol.* **232**, 508–521 (2001).
- Lundy, R. F. Jr & Contreras, R. J. Gustatory neuron types in rat geniculate ganglion. *J. Neurophysiol.* **82**, 2970–2988 (1999).
- Gilbertson, T. A., Boughter, J. D. Jr, Zhang, H. & Smith, D. V. Distribution of gustatory sensitivities in rat taste cells: whole-cell responses to apical chemical stimulation. *J. Neurosci.* **21**, 4931–4941 (2001).
- Erickson, R. P. The evolution of neural coding ideas in the chemical senses. *Physiol. Behav.* **69**, 3–13 (2000).
- Roper, S. D. Regenerative impulses in taste cells. *Science* **220**, 1311–1312 (1983).
- Avenet, P. & Lindemann, B. Patch-clamp study of isolated taste receptor cells of the frog. *J. Membr. Biol.* **97**, 223–240 (1987).
- Lindemann, B. Taste reception. *Physiol. Rev.* **76**, 719–766 (1996).
- Akabas, M. H., Dodd, J. & Al-Awqati, Q. A bitter substance induces a rise in intracellular calcium in a subpopulation of rat taste cells. *Science* **242**, 1047–1050 (1988).
- Bernhardt, S. J., Naim, M., Zehavi, U. & Lindemann, B. Changes in IP<sub>3</sub> and cytosolic Ca<sup>2+</sup> in response to sugars and non-sugar sweeteners in transduction of sweet taste in the rat. *J. Physiol.* **490**, 325–336 (1996).
- Caicedo, A. & Roper, S. D. Taste receptor cells that discriminate between bitter stimuli. *Science* **291**, 1557–1560 (2001).
- Zviman, M. M., Restrepo, D. & Teeter, J. H. Single taste stimuli elicit either increases and decreases in intracellular calcium in isolated catfish taste cells. *J. Membr. Biol.* **149**, 81–88 (1996).
- Hayashi, Y., Zviman, M. M., Brand, J. G., Teeter, J. H. & Restrepo, D. Measurement of membrane potential and [Ca<sup>2+</sup>]<sub>i</sub> in cell ensembles: application to the study of glutamate taste in mouse. *Biophys. J.* **71**, 1057–1070 (1996).
- Herness, M. S. & Sun, X. D. Characterization of chloride currents and their noradrenergic modulation in rat taste receptor cells. *J. Neurophysiol.* **82**, 260–271 (1999).
- Yamamoto, T., Nagai, T., Shimura, T. & Yasoshima, Y. Roles of chemical mediators in the taste system. *Jpn J. Pharmacol.* **76**, 325–348 (1998).
- Delay, R. J., Kinnamon, S. C. & Roper, S. D. Serotonin modulates voltage-dependent calcium currents in *Necturus* taste cells. *J. Neurophysiol.* **77**, 2515–2524 (1997).
- Herness, S. & Chen, Y. Serotonin inhibits calcium-activated K<sup>+</sup> current in rat taste receptor cells. *NeuroReport* **8**, 3257–3261 (1997).
- Caicedo, A., Kim, K. N. & Roper, S. D. Glutamate-induced cobalt uptake reveals non-NMDA receptors in rat taste cells. *J. Comp. Neurol.* **417**, 315–324 (2000).
- Lawton, D. M., Furness, D. N., Lindemann, B. & Hackney, C. M. Localization of the glutamate-aspartate transporter, GLAST, in rat taste buds. *Eur. J. Neurosci.* **12**, 3163–3171 (2000).
- Lindemann, B. Sodium taste. *Curr. Opin. Nephrol. Hypertension* **6**, 425–429 (1997).
- Heck, G. L., Mierson, S. & DeSimone, J. A. Salt taste transduction occurs through an amiloride-sensitive sodium transport pathway. *Science* **223**, 403–405 (1984).
- Canessa, C. M. *et al.* Amiloride-sensitive epithelial Na<sup>+</sup> channel is made of three homologous subunits. *Nature* **367**, 463–467 (1994).
- Kretz, O., Barbry, P., Bock, R. & Lindemann, B. Differential expression of RNA and protein of the three pore-forming subunits of the amiloride-sensitive epithelial sodium channel in taste buds of the rat. *J. Histochem. Cytochem.* **47**, 51–64 (1999).
- Lin, W., Finger, T. E., Rossier, B. C. & Kinnamon, S. C. Epithelial Na<sup>+</sup> channel subunits in rat taste cells: localization and regulation by aldosterone. *J. Comp. Neurol.* **405**, 406–420 (1999).
- Avenet, P. & Lindemann, B. Noninvasive recording of receptor cell action potentials and sustained currents from single taste buds maintained in the tongue: the response to mucosal NaCl and amiloride. *J. Membr. Biol.* **124**, 33–41 (1991).
- Smith, D. V. & Ossebaard, C. A. Amiloride suppression of the taste intensity of sodium chloride: evidence from direct magnitude scaling. *Physiol. Behav.* **57**, 773–777 (1995).
- Gilbertson, T. A., Roper, S. D. & Kinnamon, S. C. Proton currents through amiloride-sensitive Na<sup>+</sup> channels in localized hamster taste cells: enhancement by vasopressin and cAMP. *Neuron* **10**, 931–942 (1993).
- Kinnamon, S. C., Dionne, V. E. & Beam, K. G. Apical localization of K channels in taste cells provides the basis for sour taste transduction. *Proc. Natl Acad. Sci. USA* **85**, 7023–7027 (1988).
- Ugawa, S. *et al.* Receptor that leaves a sour taste in the mouth. *Nature* **395**, 555–556 (1998).
- Miyamoto, T., Fujiyama, R., Okada, Y. & Sato, T. Sour transduction involves activation of NPPB-sensitive conductance in mouse taste cells. *J. Neurophysiol.* **80**, 1852–1859 (1998).
- Miyamoto, T., Fujiyama, R., Okada, Y. & Sato, T. Acid and salt responses in mouse taste cells. *Prog. Neurobiol.* **62**, 135–157 (2000).
- Stevens, D. R. *et al.* The hyperpolarization-activated channels HCN1 and 4 mediate responses to sour stimuli. *Nature* (in the press).
- DeSimone, J. A., Callahan, E. M. & Heck, G. L. Chorda tympani taste response of rat to hydrochloric acid subject to voltage-clamped lingual receptive field. *Am. J. Physiol.* **268**, C1295–C1300 (1995).
- Stewart, R. E., Llyall, V., Feldman, G. M., Heck, G. L. & DeSimone, J. A. Acid-induced responses in hamster chorda tympani and intracellular pH tracking by taste receptor cells. *Am. J. Physiol.* **275**, C227–C238 (1998).
- Spielman, A. I., Huque, T., Whitney, G. & Brand, J. G. in *Sensory Transduction* (eds Corey, D. P. & Roper, S. D.) 307–324 (The Rockefeller University Press, New York, 1992).
- Adler, E. *et al.* A novel family of mammalian taste receptors. *Cell* **100**, 693–702 (2000).
- Matsunami, H., Montmayeur, J.-P. & Buck, L. A family of candidate taste receptors in human and mouse. *Nature* **404**, 601–604 (2000).
- Chandrasekar, J. *et al.* T2Rs function as bitter taste receptors. *Cell* **100**, 703–711 (2000).
- Dahl, M., Erickson, R. P. & Simon, S. A. Neural responses to bitter compounds in rats. *Brain Res.* **756**, 22–34 (1997).
- Naim, M., Seifert, R., Nürnberg, B., Grünbaum, L. & Schultz, G. Some taste substances are direct activators of G-proteins. *Biochem. J.* **297**, 451–454 (1994).
- Cummings, T. A. & Kinnamon, S. C. Apical K<sup>+</sup> channels in *Necturus* taste cells—modulation by intracellular factors and taste stimuli. *J. Gen. Physiol.* **99**, 591–613 (1992).
- Tsunenari, T. *et al.* A quinine-activated cationic conductance in vertebrate taste receptor cells. *J. Gen. Physiol.* **108**, 515–523 (1996).
- Spielman, A. I. *et al.* A method for isolating and patch-clamping single mammalian taste receptor cells. *Brain Res.* **503**, 326–329 (1989).
- Rosenzweig, S., Yan, W., Dasso, M. & Spielman, A. I. Possible novel mechanism for bitter taste mediated through cGMP. *J. Neurophysiol.* **81**, 1661–1665 (1999).
- Kretz, O., Bock, R. & Lindemann, B. Occurrence of nitric oxide synthase in taste buds of the rat vallate papilla. *Histochem. J.* **30**, 293–299 (1998).
- Wong, G. T., Gannon, K. S. & Margolske, R. F. Transduction of bitter and sweet taste by gustducin. *Nature* **381**, 796–800 (1996).
- Spickofsky, N. *et al.* Biochemical analysis of the transducin-phosphodiesterase interaction. *Nature Struct. Biol.* **1**, 771–781 (1994).
- Yan, W. *et al.* Bitter taste transduced by PLC- $\beta$ -dependent rise in IP<sub>3</sub> and  $\alpha$ -gustducin-dependent fall in cyclic nucleotides. *Am. J. Physiol. Cell Physiol.* **280**, C742–C751 (2001).
- Huang, L. *et al.* G $\gamma$ 13 colocalizes with gustducin in taste receptor cells and mediates IP<sub>3</sub> responses to



- bitter denatonium. *Nature Neurosci.* **2**, 1055–1062 (1999).
61. Rössler, P., Kroner, C., Freitag, J., Noé, J. & Breer, H. Identification of a phospholipase c  $\beta$  subtype in rat taste cells. *Eur. J. Cell Biol.* **77**, 253–261 (1998).
62. Rössler, P. *et al.* G protein betagamma complexes in circumvallate taste cells involved in bitter transduction. *Chem. Senses* **25**, 413–421 (2000).
63. Spielman, A. I., Huque, T., Nagai, H., Whitney, G. & Brand, J. G. Generation of inositol phosphates in bitter taste transduction. *Physiol. Behav.* **56**, 1149–1155 (1994).
64. Spielman, A. I. *et al.* Rapid kinetics of second messenger formation in bitter taste. *Am. J. Physiol. Cell Physiol.* **270**, C926–C931 (1996).
65. Clapp, T. R., Stone, L. M., Margolskee, R. F. & Kinnamon, S. C. Immunocytochemical evidence for co-expression of Type III IP<sub>3</sub> receptor with signaling components of bitter taste transduction. *BMC Neurosci.* **2**, 6 (2001).
66. Miyoshi, M. A., Abe, K. & Emori, Y. IP<sub>3</sub> receptor type 3 and PLC $\beta$ 2 are co-expressed with taste receptors T1R and T2R in rat taste bud cells. *Chem. Senses* **26**, 259–265 (2001).
67. Ogura, T. & Kinnamon, S. C. IP<sub>3</sub>-Independent release of Ca<sup>2+</sup> from intracellular stores: a novel mechanism for transduction of bitter stimuli. *J. Neurophysiol.* **82**, 2657–2666 (1999).
68. Nofre, C. & Tinti, J. M. Sweetness reception in man: the multipoint attachment theory. *Food Chem.* **56**, 263–274 (1996).
69. Ishimoto, H., Matsumoto, A. & Tanimura, T. Molecular identification of a taste receptor gene for trehalose in *Drosophila*. *Science* **289**, 116–119 (2000).
70. Ninomiya, Y., Sako, N. & Funakoshi, M. Selective effects of the dpa gene on the ability to taste D-phenylalanine in mice. *Proc. Jpn Symp. Taste Smell* **21**, 153–156 (1987).
71. Lush, I. E. The genetics of tasting in mice. VI. Saccharin, acesulfame, dulcin and sucrose. *Genet. Res.* **53**, 95–99 (1989).
72. Max, M. *et al.* *Tas1r3*, encoding a new candidate taste receptor, is allelic to the sweet responsiveness locus *Sac*. *Nature Genet.* **28**, 58–63 (2001).
73. Montmayeur, J. P., Liberles, S. D., Matsunami, H. & Buck, L. B. A candidate taste receptor gene near a sweet taste locus. *Nature Neurosci.* **4**, 492–498 (2001).
74. Kitagawa, M., Kusakabe, Y., Miura, H., Ninomiya, Y. & Hino, A. Molecular genetic identification of a candidate receptor gene for sweet taste. *Biochem. Biophys. Res. Commun.* **283**, 236–242 (2001).
75. Sainz, E., Korley, J. N., Battey, J. F. & Sullivan, S. L. Identification of a novel member of the T1R family of putative taste receptors. *J. Neurochem.* **77**, 896–903 (2001).
76. Hoon, M. A. *et al.* Putative mammalian taste receptors: a class of taste specific GPCRs with distinct topographic selectivity. *Cell* **96**, 541–551 (1999).
77. Nelson, G. *et al.* Mammalian sweet taste receptors. *Cell* **106**, 381–390 (2001).
78. Avenet, P., Hofmann, F. & Lindemann, B. Transduction in taste receptor cells requires cAMP-dependent protein kinase. *Nature* **331**, 351–354 (1988).
79. Tonosaki, K. & Funakoshi, M. Cyclic nucleotides may mediate taste transduction. *Nature* **331**, 354–356 (1988).
80. Béhé, P., DeSimone, J. A., Avenet, P. & Lindemann, B. Membrane currents in taste cells of the rat fungiform papilla: evidence for two types of Ca currents and inhibition of K currents by saccharin. *J. Gen. Physiol.* **96**, 1061–1084 (1990).
81. Striemi, B., Pace, U., Zehavi, U., Naim, M. & Lancet, D. Sweet tastants stimulate adenylate cyclase coupled to GTP-binding protein in rat tongue membranes. *Biochem. J.* **260**, 121–126 (1989).
82. Striemi, B. J., Naim, M. & Lindemann, B. Generation of cyclic AMP in taste buds of the rat circumvallate papilla in response to sucrose. *Cell. Physiol. Biochem.* **1**, 46–54 (1991).
83. Cummings, T. A., Powell, J. & Kinnamon, S. C. Sweet taste transduction in hamster taste cells: evidence for the role of cyclic nucleotides. *J. Neurophysiol.* **70**, 2326–2336 (1993).
84. Cummings, T. A., Daniels, C. & Kinnamon, S. C. Sweet taste transduction in hamster: sweeteners and cyclic nucleotides depolarize taste cells by reducing a K<sup>+</sup> current. *J. Neurophysiol.* **75**, 1256–1263 (1996).
85. Uchida, Y. & Sato, T. Changes in outward K<sup>+</sup> currents in response to two types of sweeteners in sweet taste transduction of gerbil taste cells. *Chem. Senses* **22**, 163–169 (1997).
86. Nakashima, K. & Ninomiya, Y. Transduction for sweet taste of saccharin may involve both inositol 1,4,5-trisphosphate and cAMP pathways in the fungiform taste buds in C57BL mice. *Cell. Physiol. Biochem.* **9**, 90–98 (1999).
87. Ishimaru, Y., Yasuoka, A., Asano-Miyoshi, M., Abe, K. & Emori, Y. An actin-binding protein, CAP, is expressed in a subset of rat taste bud cells. *NeuroReport* **12**, 233–235 (2001).
88. Misaka, T., Kusakabe, Y., Emori, Y., Arai, S. & Abe, K. Molecular cloning and taste bud-specific expression of a novel cyclic nucleotide-gated channel. *Ann. NY Acad. Sci.* **855**, 150–159 (1998).
89. Varkevisser, B. & Kinnamon, S. C. Sweet taste transduction in hamster: role of protein kinases. *J. Neurophysiol.* **83**, 2526–2532 (2000).
90. Krizhanovsky, V., Agamy, O. & Naim, M. Sucrose-stimulated subsecond transient increase in cGMP level in rat intact circumvallate taste bud cells. *Am. J. Physiol. Cell Physiol.* **279**, C120–C125 (2000).
91. Kawai, K., Sugimoto, K., Nakashima, K., Miura, H. & Ninomiya, Y. Leptin as a modulator of sweet taste sensitivities in mice. *Proc. Natl Acad. Sci. USA* **97**, 11044–11049 (2000).
92. Ninomiya, Y. *et al.* Leptin and sweet taste. *Vitam. Horm.* (in the press).
93. Ikeda, K. On a new seasoning. *J. Tokyo Chem. Soc.* **30**, 820–836 (1909). [In Japanese.]
94. Chaudhari, N., Landin, A. M. & Roper, S. D. A novel metabotropic glutamate receptor functions as a taste receptor. *Nature Neurosci.* **3**, 113–119 (2000).
95. Delay, E. R. *et al.* Taste preference synergy between glutamate receptor agonists and inosine monophosphate in rats. *Chem. Senses* **25**, 507–515 (2000).
96. Bigiani, A., Delay, R. J., Chaudhari, N., Kinnamon, S. C. & Roper, S. D. Responses to glutamate in rat taste cells. *J. Neurophysiol.* **77**, 3048–3059 (1997).
97. Lin, W. & Kinnamon, S. C. Physiological evidence for ionotropic and metabotropic glutamate receptors in rat taste cells. *J. Neurophysiol.* **82**, 2061–2069 (1999).
98. Brand, J. G. Receptor and transduction processes for umami taste. *J. Nutr.* **130**, 942S–945S (2000).
99. Stevens, J. C., Cruz, L. A., Hoffman, J. M. & Patterson, M. Q. Taste sensitivity and aging: high incidence of decline revealed by repeated threshold measures. *Chem. Senses* **20**, 451–459 (1995).
100. Zhu, X., Gilbert, S., Birnbaumer, M. & Birnbaumer, L. Dual signaling potential is common among G<sub>i</sub>-coupled receptors and dependent on receptor density. *Mol. Pharmacol.* **46**, 460–469 (1994).

## Acknowledgements

Owing to limitations of space, the important work of many colleagues could not be cited, for which I apologize. I thank S. C. Kinnamon and R. F. Margolskee for comments.

# Stochastic sensors inspired by biology

Hagan Bayley\*†‡ & Paul S. Cremer†‡

\*Department of Medical Biochemistry & Genetics, The Texas A&M University System Health Science Center, College Station, Texas 77843-1114, USA (e-mail: bayley@tamu.edu)

†Department of Chemistry, and ‡Center for Integrated Microchemical Systems, Texas A&M University, College Station, Texas 77843-3255, USA

**Sensory systems use a variety of membrane-bound receptors, including responsive ion channels, to discriminate between a multitude of stimuli. Here we describe how engineered membrane pores can be used to make rapid and sensitive biosensors with potential applications that range from the detection of biological warfare agents to pharmaceutical screening. Notably, use of the engineered pores in stochastic sensing, a single-molecule detection technology, reveals the identity of an analyte as well as its concentration.**

Channels and pores that respond directly to molecules or to physical stimuli are found within sensory systems. They include amiloride-sensitive  $\text{Na}^+$  channels, which detect the taste of salt, and the vanilloid receptors, which transduce both noxious stimuli, such as elevated temperature, and the sensation of hot peppers. Stochastic sensing with protein pores does not attempt to mimic the molecular devices found in nature, rather the approach is inspired by nature both in the use of protein pores and in the design of binding sites for analytes. Advances made in the past few years allow detection of small cations and anions, organic molecules, proteins and DNA.

## Stochastic sensing with single protein pores

The first example of a single-molecule experiment on an identified functional biomolecule was conducted over 30 years ago: the observation of current flow through a single ion-conducting channel formed by the peptide antibiotic gramicidin in a planar lipid bilayer<sup>1</sup>. Soon afterwards, it was shown that single-channel currents can be modulated by molecules, known as channel blockers, that bind reversibly within the lumen of the channel. This finding forms the basis for stochastic sensing with engineered pores<sup>2,3</sup>.

To understand stochastic sensing, the advantages of detection at the single-molecule level must be appreciated (Box 1). In the simplest case, the sensor element has two states — occupied (by analyte) and unoccupied — and a different output is associated with each. As well as revealing analyte concentration, stochastic sensing provides structure-specific information about an analyte, which is often sufficient for its identification.

The pores used for stochastic sensing are based on staphylococcal  $\alpha$ -haemolysin ( $\alpha\text{HL}$ ). The pore formed by wild-type  $\alpha\text{HL}$  consists of seven identical subunits arranged around a central axis. The transmembrane part of the lumen is a  $\beta$ -barrel with two antiparallel strands contributed by each subunit. The extramembraneous domain contains a large cavity that houses the transmembrane domain during the assembly process, but which is available for engineering in the assembled pore. Although channels and pores have been used for some time for analyte detection<sup>4,5</sup>, protein engineering and single-molecule detection greatly enhance the potential of the technology<sup>6,7</sup>. The three-dimensional structure of the  $\alpha\text{HL}$  pore is known at high resolution and, importantly,  $\beta$ -barrels are especially open to remodelling<sup>8</sup>. The large

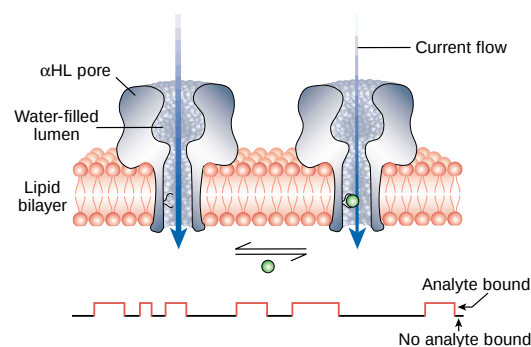
single-channel conductance ( $\sim 1$  nS in 1 M KCl) makes single-channel recording straightforward.

## Engineering pores for a wide variety of analytes

Binding sites for analytes have been placed within the lumen of the  $\alpha\text{HL}$  pore by a variety of means, and pores have also

### Box 1

#### Stochastic sensing with engineered pores



A single engineered pore is placed in a planar lipid bilayer. In an applied potential, a current flows through the pore carried by the ions in salt solutions that bath both side of the bilayer. The pore contains a binding site for an analyte, represented by the green ball in the figure. Each time the analyte binds to the pore the current is modulated as illustrated in the trace. Hence, this technique monitors individual binding events. The frequency of occurrence of the events reveals the concentration of the analyte, whereas the current signature (the mean duration and amplitude of the events) reveals its identity.

In the case of a simple equilibrium,  $\tau_{\text{off}} = 1/k_{\text{off}}$ , where  $\tau_{\text{off}}$  is the mean dwell time of the analyte and  $k_{\text{off}}$  is the dissociation rate constant, and  $\tau_{\text{on}} = 1/k_{\text{on}}[A]$ , where  $\tau_{\text{on}}$  is the mean time between binding events,  $k_{\text{on}}$  is the association rate constant and  $[A]$  is the analyte concentration

In this case, stochastic sensing provides useful kinetic data that are difficult to obtain by other techniques. In other cases, the binding kinetics are more complex, but this can be useful by providing highly distinctive current signatures. In practice, it is not necessary to understand the details of the kinetic behaviour to allow the identification and quantification of analytes.

been engineered so that binding of an analyte in the external solution produces a reorganization within the lumen. Consequently, the current passing through a single pore responds to each binding event. Although this kind of stochastic sensing does not occur in nature, the binding sites used are based upon natural examples.

#### Detection of ionic species

Engineered versions of the  $\alpha$ HL pore have been used to detect divalent metal ions ( $M(II)$ )<sup>6,9</sup>. In the best studied case, residues with side chains projecting into the lumen were replaced with four histidines to form the mutant subunit 4H. This was incorporated into a heteromeric pore comprising six wild-type subunits and one 4H subunit (WT<sub>6</sub>4H<sub>1</sub>). The resulting pore contained a single  $M(II)$ -binding site resembling that of carbonic anhydrase. Single-channel recordings showed that WT<sub>6</sub>4H<sub>1</sub> responds to nanomolar concentrations of Zn(II) (Fig. 1a). Single-molecule detection also enables mixtures of ions to be analysed — more than one analyte can bind to the  $M(II)$ -binding site of WT<sub>6</sub>4H<sub>1</sub>, but because binding is mutually exclusive, only one  $M(II)$  is bound and detected at any given moment, and the identity of each  $M(II)$  is revealed by its characteristic signature (Fig. 1a). The approach can be applied to other charged analytes. For example,  $\alpha$ HL pores that detect anions have been prepared recently by replacement of residues within the  $\beta$ -barrel by mutagenesis (S. Cheley and L. Gu, unpublished results).

#### Comparison of multianalyte detection by the olfactory system

The mammalian olfactory system also uses 'crossreactive' receptors, each species of which binds a different set of analytes. The receptors differ in their affinities for analytes, such that the sensor array of the olfactory epithelium, which contains hundreds of receptor species

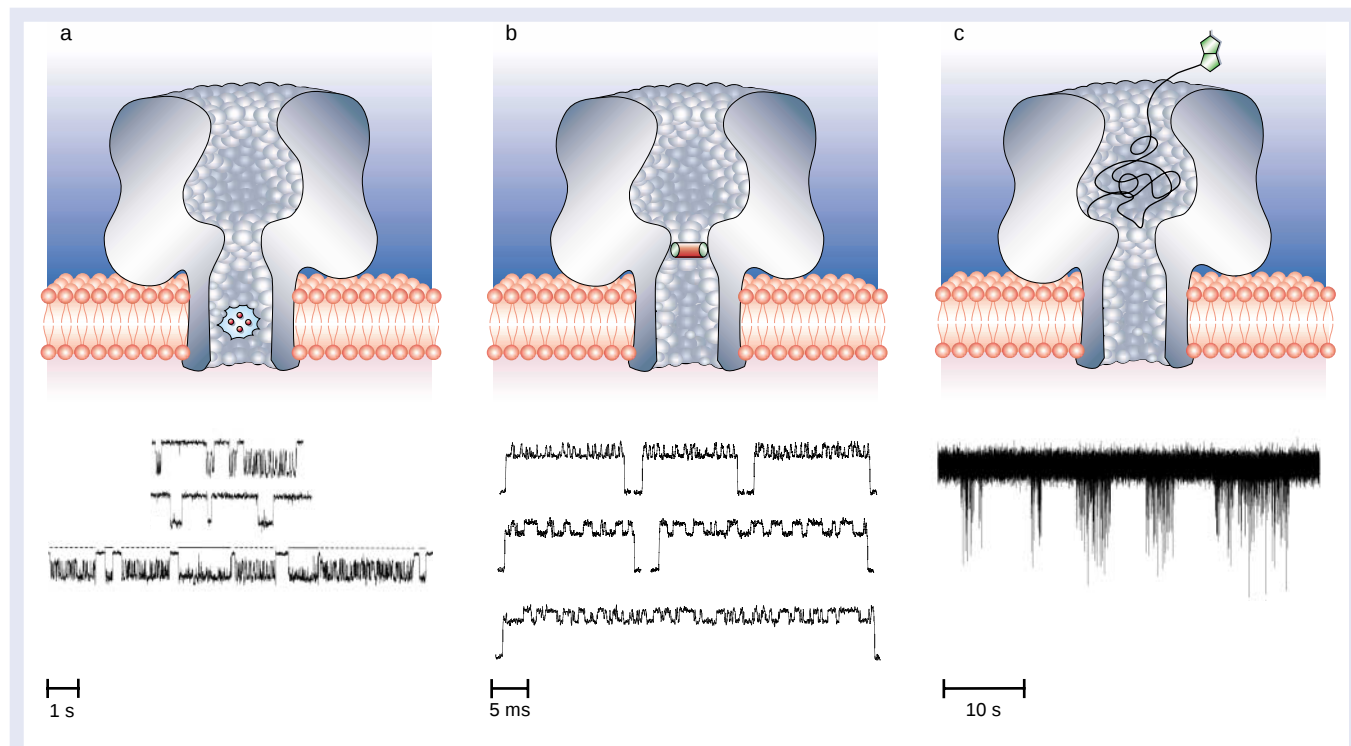
each confined to a distinct set of cells, is capable of identifying thousands of individual odours. By contrast, analytes are identified in stochastic sensing from their current signatures. Combining the two ways of distinguishing signals from crossreactive receptors, by using arrays of stochastic sensing elements, could provide a powerful technology for analyte detection.

Because sensor elements capable of binding a variety of analytes can be used in stochastic sensing, the required genetic engineering is simplified. By contrast, specific receptor sites are necessary where protein-based biosensors are used with outputs derived from ensemble properties. The application of crossreactive sensor elements might be extended to the notion of sensing groups of analytes. Stochastic sensors that reliably detect a class of molecules such as organophosphorus agents would be of considerable interest.

Families of genes encode receptors in the olfactory system. By comparison, large numbers of related responsive pores can be generated in a combinatorial approach. In the case of  $M(II)$ , residues with side chains capable of coordination, such as histidine, aspartate, glutamate and cysteine, can be patterned in hundreds of ways on the interior surface of a  $\beta$ -barrel, and wild-type and mutant subunits can be arranged within the heptamer in numerous combinations and permutations<sup>6</sup>.

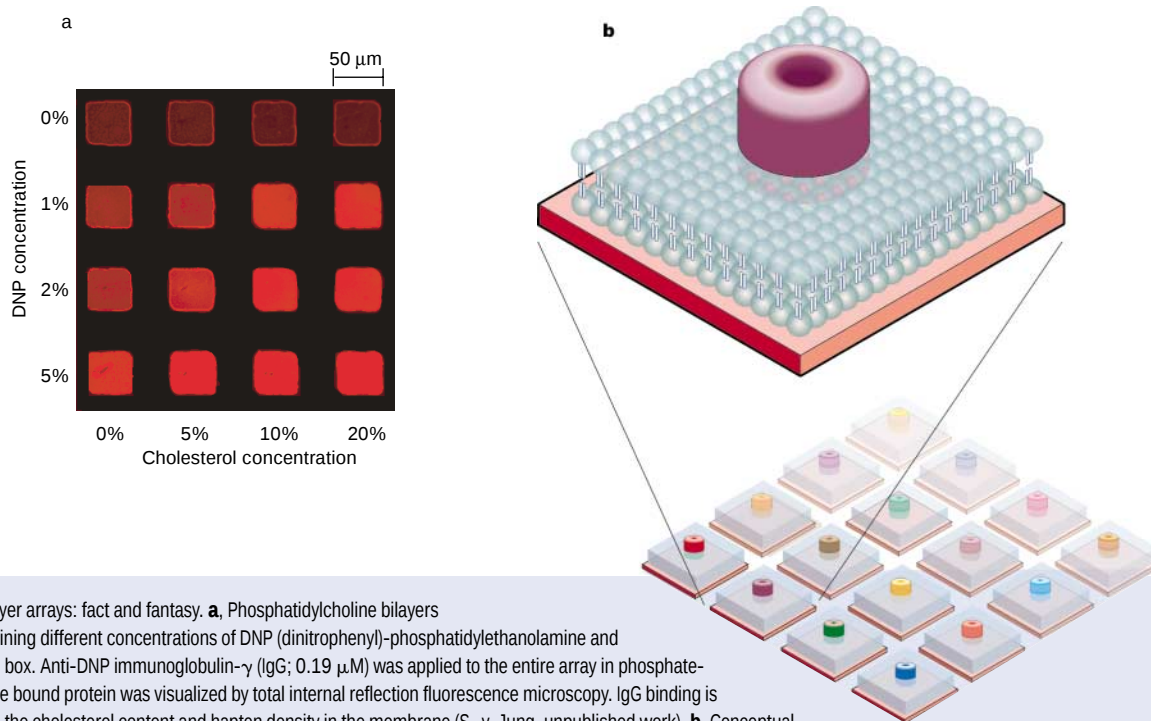
#### Detection of organic molecules

A similar approach might be applied to the analysis of organic molecules<sup>10</sup>. Although it is more difficult to design binding sites in this case, it is possible to place non-covalent adapters in the lumen of the  $\alpha$ HL pore for periods sufficiently long to observe host-guest interactions<sup>10</sup>. For example, when  $\beta$ -cyclodextrin is lodged in the pore, it remains capable of binding the same organic molecules that it binds



**Figure 1** Detection of a variety of analytes by stochastic sensing. **a**, Metal ions<sup>6,9</sup>. The sensor element is an  $\alpha$ HL pore with a genetically engineered binding site for divalent metal cations within the lumen. The site comprises four histidines on one of the seven subunits. Various metal ions are coordinated by the engineered side chains, which project into the transmembrane  $\beta$ -barrel. Each metal ion binds with different kinetics and gives slightly different extents of channel block, so that analysis of extended current traces allows the identification and quantification of metal ions in mixtures. Upper trace, Zn<sup>2+</sup> alone; middle trace, Co<sup>2+</sup>; lower trace, mixture of Zn<sup>2+</sup> (0.23  $\mu$ M) and Co<sup>2+</sup> (4.7  $\mu$ M). **b**, Organic molecules<sup>10</sup>. The pore contains a non-covalent  $\beta$ -cyclodextrin adapter, which is capable of carrying out host-guest chemistry while lodged in the lumen. Upper trace, promethazine; middle trace, imipramine; lower trace, mixture of promethazine (100  $\mu$ M) and imipramine (100  $\mu$ M). **c**, Proteins<sup>21</sup>. The pore contains a single poly(ethylene glycol) chain of relative molecular mass 3,400. One end of the chain is attached covalently within the cavity of the  $\alpha$ HL pore, the other end carries a ligand for the target protein. The trace shows the response of a pore carrying a biotin ligand to a mutant streptavidin (14.5 nM) of weakened affinity.





**Figure 2** Lipid bilayer arrays: fact and fantasy. **a**, Phosphatidylcholine bilayers ( $4 \times 4$  array) containing different concentrations of DNP (dinitrophenyl)-phosphatidylethanolamine and cholesterol in each box. Anti-DNP immunoglobulin- $\gamma$  (IgG;  $0.19 \mu\text{M}$ ) was applied to the entire array in phosphate-buffered saline. The bound protein was visualized by total internal reflection fluorescence microscopy. IgG binding is dependent on both the cholesterol content and hapten density in the membrane (S.-y. Jung, unpublished work). **b**, Conceptual rendition of a stochastic sensor array. Each box contains a different, single, engineered  $\alpha\text{HL}$  pore. The bilayers are arrayed over a system of electrodes.

when free in solution; these binding events cause current fluctuations that permit the quantification and identification of the molecules<sup>10</sup> (Fig. 1b). Multiple molecules can be detected concurrently with the same cyclodextrin<sup>10</sup>, and the system is highly flexible and combinatorial. Additional adapters can be used (for example, cyclic peptides<sup>11</sup>) and the pore can be engineered to accommodate one, or even two, adapters for far longer periods than the wild-type protein<sup>12</sup>.

#### Detection of macromolecules

Single-stranded RNA or DNA molecules can pass through the wild-type  $\alpha\text{HL}$  pore in an elongated conformation<sup>13–17</sup>. The transit time and extent of current block reveal information about the length of the nucleic acid and its base composition<sup>13,15,16</sup>. In certain cases, in a step beyond what would be offered by a molecular Coulter counter<sup>3,18</sup>, the magnitude of the conductance changes within each blocking event communicates additional details about nucleic acid structure, such as the composition of stretches of  $\sim 50$  bases<sup>14</sup> or the presence of mismatches in DNA hairpins<sup>19</sup>.

Protein engineering of the  $\alpha\text{HL}$  pore greatly increases the possibilities for stochastic sensing of macromolecules. When single-stranded oligonucleotides are covalently attached within the large cavity of the pore, sequence-specific duplex formation can be detected<sup>20</sup>. But protein analytes are too large to fit inside the pore and a means to couple external binding with events in the lumen is required. This was accomplished by anchoring one end of a polymer covalently inside the lumen and attaching biotin to the other end. The motion of the polymer within the pore was altered when streptavidin or an antibody captured the biotin in the external medium, producing a change in the single-channel conductance (Fig. 1c)<sup>21</sup>. The approach might be modified for use with a wide variety of protein ligands.

#### Advantages of stochastic sensing

Stochastic sensing is highly sensitive, the response is rapid and reversible (allowing real-time monitoring of analytes) and the

dynamic range is wide. In addition to detecting both the concentration and identity of an analyte, several analytes can be quantified concurrently by a single sensor element. The sensor element need not be highly selective, as each analyte produces a characteristic signature. This has the advantage of easing the demands for protein engineering, as well as preventing confusion arising from signals from analytes that are structurally similar to the target molecule. Fouling of the sensor element cannot give a false reading, as the signal would not be characteristic of an analyte. Further advantages include no loss of signal-to-noise at low analyte concentrations (permitting the detection of rare events), a 'digital' output for facile electronic interfacing, self-calibration, reagentless operation, and the potential for nanoscale miniaturization.

#### Fabricating durable and practical devices

##### Improved apertures

One disadvantage of the current generation of stochastic sensors is their lack of durability, which confines their use to the laboratory. Several crucial advances will be required to make a practicable device. Present platforms use  $\alpha\text{HL}$  in lipid bilayers suspended across apertures in a polymer film such as Teflon. The apertures are typically a few hundred microns in diameter, and the bilayers often rupture after a few hours of use. Several strategies could be implemented to remedy this problem. Nanoscale apertures (10 nm to  $1 \mu\text{m}$ ) might limit the magnitude of bilayer undulations, which probably are a source of breakage<sup>22</sup>. The current revolution in lithography should make it possible to limit the hole size to an area slightly bigger than the protein pore itself. This would not only increase membrane durability, but also allow apertures to be designed that accommodate only a single protein pore, thereby facilitating the manufacturability of stochastic devices. Recent work with larger microfabricated apertures in silicon-based substrates shows promise<sup>23</sup>. The smallest apertures reported so far that have been covered with bilayers are  $\sim 30$  nm in diameter, but more work needs to be done to lower the capacitive noise in these systems<sup>24</sup>. It might even be possible to dispense with a

bilayer and use a hole into which engineered pores fit snugly. For example, gold-coated nanotubes can be created by plating etched particle tracks in polycarbonate filters. In this case, the surface of the nanotube and the pore might be tailored for fit and compatibility<sup>3</sup>.

#### Bilayers on solid supports

Another approach to rugged sensor design could involve the incorporation of  $\alpha$ HL into bilayers on solid supports<sup>25,26</sup>. When the support is a conducting substrate, an alternating potential can be applied to allow impedance measurements. But impedance spectroscopy of tethered supported bilayers treated with  $\alpha$ HL suggests that the protein does not fully penetrate the membrane, so improvements will be needed to obtain fully conducting pores<sup>27</sup>. Although sensors based on supported bilayers have been made with several other channel-forming peptides and proteins<sup>28–31</sup>, single-channel recordings — the prerequisite for stochastic sensing — have not yet been obtained. In this regard, a major challenge is the production of tightly sealed bilayers to reduce leakage currents to the levels found in unsupported membranes. Most likely, this will require the reduction of substrate roughness and the elimination of edge effects. Through a combination of improved substrates, new membrane chemistry and faster electronics, single-channel measurements with supported bilayers at millisecond time resolution should be achieved in the near future<sup>32</sup>.

#### Bilayers on porous supports

Porous supports might provide a middle ground between apertures and solid supports. They would reinforce the bilayer and at the same time provide access for aqueous ions to the entrance of a pore. Possibilities include nanochannel glass arrays<sup>33</sup> and coated polycarbonate membranes<sup>34</sup>. Bacterial S layers are especially promising as supports. S layers are porous, two-dimensional crystals of a single protein that envelope many species of bacteria. Single-channel currents have been observed from  $\alpha$ HL incorporated into bilayers supported by S layers<sup>35,36</sup>. Peptide antibiotics and an anion channel have been incorporated into bilayers on agarose-coated glass and single-channel currents detected<sup>37</sup>.

#### Sensor arrays

The use of conducting planar supports not only has the advantage of durability, but also provides a convenient environment for patterning arrays of membranes. Ultimately, individual electrodes in the array would act as transducers for each sensing element. Lithographic patterning of very small membrane sectors may also help eliminate problems with current leakage — as the size of the bilayer elements is shrunk, a larger fraction of them will be positioned over defect-free portions of the surface. Significant progress has been made on the fabrication of supported bilayer arrays<sup>38</sup>. For example, small arrays have been made in which individual 'boxes' have different lipid compositions (Fig. 2a)<sup>39</sup> and a means by which each box can be contacted by a different aqueous phase has been devised<sup>40</sup>. Nevertheless, considerably more progress will be needed to produce arrays of single pores that are individually electrically addressed (Fig. 2b).

In the future, the platforms described above will be combined with microfluidics to create lab-on-a-chip technology<sup>41</sup>. Fabrication of linear microfluidic arrays in which both the lipid bilayers and aqueous media can be varied has already been achieved<sup>42</sup>. Arrays of sensor elements will be patterned into a substrate connected to a system of microchannels through which complex mixtures of analytes can be delivered. The array would constitute a sophisticated electronic 'nose' or 'tongue' consisting of single-molecule sensors from which a stochastic response pattern would be read. Data processing approaches such as computational neural networks have been used for other types of sensor arrays<sup>43</sup> and are likely to be effective with stochastic systems. The detection devices would be powerful enough to tackle broad classes of chemical and biological analytes with millisecond time resolution and ultra-low concentra-

tion limits. Marrying the emerging technology of stochastic sensing to microfluidics affords the additional benefits of high throughput and very small sample volumes. Of course, in this case, miniaturization will not be limited by the size of the sensor elements as a single  $\alpha$ HL pore is only ~10 nm in diameter.

#### Prospects and alternative implementation

Stochastic sensing platforms should find use in diverse applications from the detection of biological warfare agents to the testing of groundwater for heavy-metal contaminants. One of the most powerful potential applications may be in pharmaceutical screening based on proteomics, where fast time response, trace analyte detection and high throughput are required<sup>44</sup>. A significant characteristic of the technique is that the analytes, whether they be small or large molecules, need not be tagged or labelled for detection<sup>44</sup>. Also of note is that stochastic sensing reveals quantitative kinetic information (on and off rates) about the interaction of an analyte with its binding site. Indeed, stochastic sensing can operate under equilibrium conditions with very fast time resolution (hundreds of microseconds) without the complications arising from crowded chip surfaces such as unwanted polyvalency and rebinding. By using single  $\alpha$ HL pores decorated with multiple ligands in a predetermined pattern, it should be possible to examine directly the multivalent attachment of proteins<sup>45</sup> in a far cleaner way than is presently possible. And stochastic sensing has exciting potential applications in basic science. For example, analytes might be detected inside cells by a variation of patch-clamping<sup>7,46</sup> in which a responsive pore is incorporated into a patch pipette that is inserted into the cell under examination.

Although engineered protein pores hold sway for the present, stochastic sensing might be executed in other ways. Nanoscale alternatives to protein-based pores are being sought. Tubules with diameters of less than 2 nm can be obtained by electroplating pores of larger diameter<sup>47</sup>, although they have not yet been isolated as single entities. However, single carbon nanotubes, 150 nm in internal diameter, have been characterized<sup>48</sup>, and further reduction in diameter should be possible. In a recent breakthrough, apertures as small as 2 nm in diameter have been made in silicon nitride, and the transport of double-stranded DNA through a 5-nm version has been detected<sup>49</sup>. Eventually, it might be possible to manipulate the chemistry of these entities as readily as proteins.

Alternatives to ion channel-based stochastic sensing include single-molecule fluorescence<sup>50,51</sup> and force detection<sup>52,53</sup>. If practicable implementation can be found, data from fluorescence-based stochastic sensing would be handled in the same way as that obtained through electrical measurements<sup>51</sup>. Exploiting these newer technologies may open up a wider field of potential analyte targets, because not all binding sites are readily incorporated into protein pores.

The acquired wisdom of nature has made several contributions to stochastic sensing: the use of protein pores, the design of binding sites, the sensing of unmodified analytes and, potentially, the use of sensor arrays. Yet, like the best of biomimetics, the development of stochastic sensing has not blindly followed nature but it has evolved in its own way<sup>54</sup>. With more exploration to be done, the final embodiment of this powerful technology cannot yet be foreseen. □

- Hladky, S. B. & Haydon, D. A. Discreteness of conductance change in biomolecular lipid membranes in the presence of certain antibiotics. *Nature* **225**, 451–453 (1970).
- Bayley, H., Braha, O. & Gu, L.-Q. Stochastic sensing with protein pores. *Adv. Mater.* **12**, 139–142 (2000).
- Bayley, H. & Martin, C. R. Resistive-pulse sensing: from microbes to molecules. *Chem. Rev.* **100**, 2575–2594 (2000).
- Hume, R. I., Role, L. W. & Fischbach, G. D. Acetylcholine release from growth cones detected with patches of acetylcholine receptor-rich membranes. *Nature* **305**, 632–634 (1983).
- Young, S. H. & Poo, M. M. Spontaneous release of transmitter from growth cones of embryonic cells. *Nature* **305**, 634–637 (1983).
- Braha, O. et al. Designed protein pores as components for biosensors. *Chem. Biol.* **4**, 497–505 (1997).
- Trivedi, B. & Kramer, R. H. Real-time patch clamp detection of intracellular cGMP reveals long-term suppression of responses to NO and muscarinic agents. *Neuron* **21**, 895–906 (1998).
- Cheley, S., Braha, O., Lu, X., Conlan, S. & Bayley, H. A functional protein pore with a "retro" transmembrane domain. *Protein Sci.* **8**, 1257–1267 (1999).
- Braha, O. et al. Simultaneous stochastic sensing of divalent metal ions. *Nature Biotechnol.* **17**, 1005–1007 (2000).

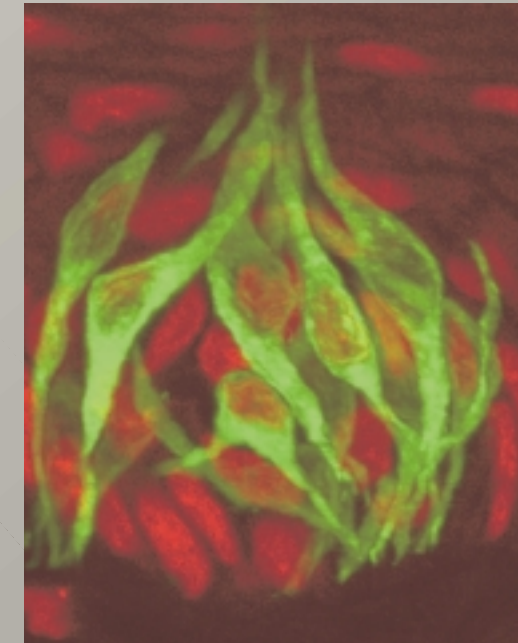
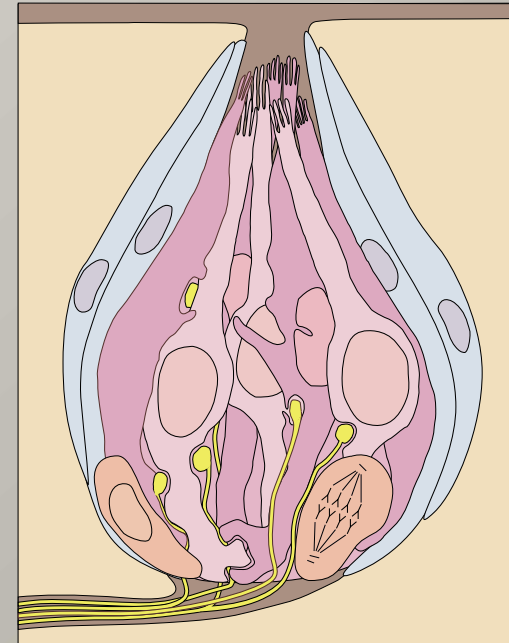
10. Gu, L.-Q., Braha, O., Conlan, S., Cheley, S. & Bayley, H. Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature* **398**, 686–690 (1999).
11. Sanchez-Quesada, J., Ghadiri, M. R., Bayley, H. & Braha, O. Cyclic peptides as molecular adapters for a pore-forming protein. *J. Am. Chem. Soc.* **122**, 11758–11766 (2000).
12. Gu, L.-Q., Cheley, S. & Bayley, H. Capture of a single molecule in a nanocavity. *Science* **291**, 636–640 (2001).
13. Kasianowicz, J. J., Brandin, E., Branton, D. & Deamer, D. W. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl Acad. Sci. USA* **93**, 13770–13773 (1996).
14. Akeson, M., Branton, D., Kasianowicz, J. J., Brandin, E. & Deamer, D. W. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.* **77**, 3227–3233 (1999).
15. Meller, A., Nivon, L., Brandin, E., Golovchenko, J. & Branton, D. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl Acad. Sci. USA* **97**, 1079–1084 (2000).
16. Henrickson, S. E., Misakian, M., Robertson, B. & Kasianowicz, J. J. Driven DNA transport into an asymmetric nanometer-scale pore. *Phys. Rev. Lett.* **85**, 3057–3060 (2000).
17. Deamer, D. W. & Akeson, M. Nanopores and nucleic acids: prospects for ultrarapid sequencing. *Trends Biotechnol.* **18**, 147–151 (2000).
18. Bezrukov, S. M. Ion channels as molecular Coulter counters to probe metabolite transport. *J. Membr. Biol.* **174**, 1–13 (2000).
19. Vercoutere, W. *et al.* Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nature Biotechnol.* **19**, 248–252 (2001).
20. Howorka, S., Cheley, S. & Bayley, H. Sequence-specific detection of individual DNA strands using engineered nanopores. *Nature Biotechnol.* **19**, 636–639 (2001).
21. Movileanu, L., Howorka, S., Braha, O. & Bayley, H. Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nature Biotechnol.* **18**, 1091–1095 (2000).
22. Hirn, R., Bayerl, T. M., Rädler, J. O. & Sackmann, E. Collective membrane motions of high and low amplitude, studied by dynamic light scattering and micro-interferometry. *Faraday Disc.* **111**, 17–30 (1998).
23. Schmidt, C., Mayer, M. & Vogel, H. A chip-based biosensor for the functional analysis of single ion channels. *Angew. Chem. Int. Edn Engl.* **39**, 3137–3140 (2000).
24. McGeoch, J. E. M., McGeoch, M. W., Carter, D. J. D., Shuman, R. F. & Guidotti, G. Biological-to-electronic interface with pores of ATP synthase subunit c in silicon nitride barrier. *Med. Biol. Eng. Comput.* **38**, 113–119 (2000).
25. Sackmann, E. Supported membranes: scientific and practical applications. *Science* **271**, 43–48 (1996).
26. Sackmann, E. & Tanaka, M. Supported membranes on soft polymer cushions: fabrication, characterization and applications. *Trends Biotechnol.* **18**, 58–64 (2000).
27. Glazier, S. A. *et al.* Reconstitution of the pore-forming toxin  $\alpha$ -hemolysin in phospholipid/18-octadecyl-1-thiahexa(ethylene oxide) and phospholipid/n-octadecanethiol supported bilayer membranes. *Langmuir* **16**, 10428–10435 (2000).
28. Cornell, B. A. *et al.* A biosensor that uses ion-channel switches. *Nature* **387**, 580–583 (1997).
29. Stora, T., Lakey, J. H. & Vogel, H. Ion-channel gating in transmembrane receptor proteins: functional activity in tethered lipid membranes. *Angew. Chem. Int. Edn Engl.* **38**, 389–392 (1999).
30. Lucas, S. W. & Harding, M. M. Detection of DNA via an ion channel switch biosensor. *Analyt. Biochem.* **282**, 70–79 (2000).
31. Straub, B., Meyer, E. & Fromherz, P. Recombinant maxi-K channels on transistor, a prototype of ionic-electronic interfacing. *Nature Biotechnol.* **19**, 121–124 (2001).
32. Wiegand, G., Neumaier, K. R. & Sackmann, E. Fast impedance spectroscopy: general aspects and performance study for single ion channel measurements. *Rev. Sci. Instrum.* **71**, 2309–2320 (2000).
33. Tonucci, R. J., Justus, B. L., Campillo, A. J. & Ford, C. E. Nanochannel array glass. *Science* **258**, 783–785 (1992).
34. Jirage, K. B., Hulteen, J. C. & Martin, C. R. Nanotube-based molecular-filtration membranes. *Science* **278**, 655–658 (1997).
35. Schuster, B., Pum, D., Braha, O., Bayley, H. & Sleytr, U. B. Self-assembled  $\alpha$ -hemolysin pores in an S-layer supported lipid bilayer. *Biochim. Biophys. Acta* **1370**, 280–288 (1998).
36. Schuster, B. *et al.* S-layer ultrafiltration membranes: a new support for stabilizing functionalized lipid membranes. *Langmuir* **17**, 499–503 (2001).
37. Ide, T. & Yanagida, T. An artificial lipid bilayer formed on an agarose-coated glass for simultaneous electrical and optical measurement of single ion channels. *Biochem. Biophys. Res. Commun.* **265**, 595–599 (1999).
38. Boxer, S. G. Molecular transport and organization in supported bilayer membranes. *Curr. Opin. Chem. Biol.* **4**, 704–709 (2000).
39. Cremer, P. S. & Yang, T. Creating spatially addressed arrays of planar supported fluid phospholipid membranes. *J. Am. Chem. Soc.* **121**, 8130–8131 (1999).
40. Yang, T., Simanek, E. E. & Cremer, P. S. Creating addressable aqueous microcompartments above solid supported bilayers using lithographically patterned PDMS molds. *Analyt. Chem.* **72**, 2587–2589 (2000).
41. Figeys, D. & Pinto, D. Lab-on-a-chip: a revolution in biological and medical sciences. *Analyt. Chem.* **72**, 330A–335A (2000).
42. Yang, T., Jung, S., Mao, H. & Cremer, P. S. Fabrication of phospholipid bilayer-coated microchannels for on-chip immunoassays. *Analyt. Chem.* **73**, 165–169 (2001).
43. Dickinson, T. A., White, J., Kauer, J. S. & Walt, D. R. A chemical-detecting system based on a cross-reactive optical sensor array. *Nature* **382**, 687–700 (1996).
44. Kodadek, T. Protein microarrays: prospects and problems. *Chem. Biol.* **8**, 105–115 (2001).
45. Borman, S. Multivalency: strength in numbers. *Chem. Eng. News* 48–53 (9 October 2000).
46. Kramer, R. H. Patch cramming: monitoring intracellular messengers in intact cells with membrane patches containing detector ion channels. *Neuron* **4**, 335–341 (1990).
47. Hulteen, J. C., Jirage, K. B. & Martin, C. R. Introducing chemical transport selectivity into gold nanotube membranes. *J. Am. Chem. Soc.* **120**, 6603–6604 (1998).
48. Sun, L. & Crooks, R. M. Single carbon nanotube membranes: a well-defined model for studying mass transport through nanoporous materials. *J. Am. Chem. Soc.* **122**, 12340–12345 (2000).
49. Li, J. *et al.* Ion beam sculpting on the nanoscale. *Nature* **412**, 166–169 (2001).
50. Weiss, S. Fluorescence spectroscopy of single biomolecules. *Science* **283**, 1676–1683 (1999).
51. Xie, X. S. & Lu, H. P. Single-molecule enzymology. *J. Biol. Chem.* **274**, 15967–15970 (1999).
52. Mehta, A. D., Rief, M., Spudich, J. A., Smith, D. A. & Simmons, R. M. Single-molecule biomechanics with optical methods. *Science* **283**, 1689–1695 (1999).
53. Fisher, T. E. *et al.* Single molecule force spectroscopy of modular proteins in the nervous system. *Neuron* **27**, 435–446 (2000).
54. Ball, P. Life's lessons in design. *Nature* **409**, 413–416 (2001).

## Acknowledgements

The investigation of stochastic sensors in the Bayley laboratory has been supported by DARPA, DOE, NIH, ONR and the Texas Advanced Technology Program. The development of membrane arrays in the Cremer laboratory has been supported by ONR, ARO, ACS-PRF, 3M Corporation and the Texas Advanced Technology Program. We thank the members of our laboratories for their energetic pursuit of these projects, and D. Deamer and J. Kauer for comments on the manuscript.



# SENSATION!



*Left:* An illustration depicting a rat vallate taste bud. The taste bud contains spindle-shaped receptor cells that send processes toward the lingual surface into the taste pore, the site of interactions with taste substances. *Right:* Immunolocalization of gustducin within the spindle-shaped cells indicated by green immunofluorescence. Gustducin is a G-protein that is involved in bitter-and-sweet-taste signal transduction. The red immunofluorescent nuclear marker indicates that gustducin is expressed only in a subset of the receptor cells. Photographs provided by David V. Smith, University of Maryland School of Medicine.

## MOLECULAR SENSING

### MOLECULAR SENSING IN HEARING AND BALANCE

More than 28 million Americans are deaf or hearing impaired, making this the third most common chronic condition affecting the elderly. As many as three of every 1,000 infants are born with a hearing impairment. Age related hearing loss (presbycusis) affects 30 to 35% of the population over the age of 65. Middle ear infections (otitis media) rank first among the diagnoses requiring a visit to the doctor or emergency room in young children. Hearing impairments occurring later in life are frequently caused by environmental factors, such as ototoxic drugs and exposure to noise. Unfortunately most hearing loss is permanent. Hearing plays a vitally important role in the quality of life. The loss of hearing can have severe socioeconomic impact upon the individual.

continued



National  
Institute on  
Deafness and  
Other  
Communication  
Disorders



National  
Institute of  
Dental and  
Craniofacial  
Research



National  
Institute of  
Neurological  
Disorders  
and Stroke



National Eye  
Institute

# MOLECULAR SENSING



National  
Institute on  
Deafness and  
Other  
Communication  
Disorders

## *continued from the inside cover*

Vestibular and/or balance problems are reported in about 9% of the population who are 65 years of age or older. Fall-related injuries such as hip fracture are a leading cause of death and disability in the elderly population. Many of the hip fractures sustained in elderly individuals by falling are related to balance disorders. The failure to diagnose and treat even benign causes of dizziness and imbalance has serious health-care implications.

Disorders of balance and spatial orientation constitute major health problems, adversely limiting an individual's sense of independence, vocational opportunities and quality of life.

A genetic approach has been useful for identifying key components of the auditory and vestibular signal transduction pathways. While less is known about the genetic and molecular mechanisms controlling the development, function and aging of sensory apparatus of the vestibular system, remarkably, 19 new genes involved in human non-syndromic deafness have been identified over the past four years. Complementary genetic studies in mice have identified over 90 different genes that affect cochlear development or function critical to hearing. A range of molecules involved in hearing have been identified, such as ion channels and pumps, gap junctions, motor molecules, transcription factors, extracellular matrix components and mitochondrial proteins. Both extramural and intramural NIDCD scientists continue to collaborate with investigators around the world to identify, isolate and study the genes that encode the diverse molecules that comprise the auditory sensing apparatus.

Recent studies in hearing and balance have led to major advancements in our understanding of the signal transduction processes in the auditory and vestibular systems. The auditory process involves a multitude of mechanical and biochemical actions that translate acoustic vibrations into the ultimate perception of sound. The sensory cells of hearing, the hair cells, are located within the cochlea in the organ of Corti. Hair cell stereocilia can be viewed as true molecular sensor organelles. Deflection of the stereocilia induces selective signal transduction processes with subsequent activation of auditory nerve fibers. The exact sequence of events is not well understood, with possible regulatory events in both the stereocilia and hair cells. Regulation at the level of the hair cell may involve the membrane protein prestin, a novel protein with shared homology to another protein, pendrin, a member of a family of sulphate/anion transport proteins. Interestingly, a mutation in the gene encoding pendrin is the cause of the genetically inherited deafness observed in Pendred's syndrome.

Human balance control is a complex, multimodal sensorimotor function, involving several sensory input modalities, and motor output systems. The primary function of the vestibular system is to provide these various functional components of the nervous system with information on the orientation and motion of the head to maintain balance and to stabilize the retinal image. The identification and therapeutic relief of vertigo and other vestibular afflictions have been slow in coming. One potential treatment identified by NIDCD-funded research suggests symptomatic relief can be achieved by alteration of the vestibular semi-circular canals.

## MOLECULAR SENSING IN THE CHEMICAL SENSES

Three million Americans experience persistent difficulty with their senses of smell and/or taste. The higher incidence of the loss of smell, whether partial (hyposmia) or complete (anosmia), is accompanied by a loss of flavor perception and enjoyment of food. In elderly individuals, a loss of smell can have a significant impact upon health and well being. While altered taste sensitivity does play a role in nutrition and weight control, this taste alteration (hypogeusia) or complete loss (ageusia) is relatively uncommon, and complaints about food blandness are usually due to olfactory problems. Unfortunately, there are no treatments for these chemosensory disorders, and scientists are aggressively seeking the underlying mechanisms that trigger the loss or diminution of these important sensory systems.

The chemical senses possess an incredible repertoire of receptors. Both NIDCD-funded intramural and extramural laboratories have contributed to the recent advances in our understanding of the signal transduction processes that underlie the perception of odors and tastants. Two landmark discoveries occurred in the field of olfaction in the early 1990s when separate families of genes were identified that encode the G-protein coupled receptors in the neurons of the olfactory and vomeronasal epithelia. These successful research strategies were later adapted to the gustatory system, culminating with the recent discovery of taste receptor genes encoding the G-protein coupled receptors activated by bitter- and sweet-tasting substances. Further, the identification of genes important to the taste and smell processes have been important in understanding how these systems function. Genes identified for bitter-tasting compounds, and a putative sweet taste receptor gene, *Tas1r3*, near the *Sac* locus on mouse chromosome 4, a determinant of sweet preference, may eventually influence the treatment of diabetes and obesity, which affect millions of Americans.

The NIDCD supports and conducts research and research training in the normal and disordered processes of hearing, balance, smell, taste, voice, speech, and language. NIDCD supports both basic and clinical biomedical research in laboratories across the country through a peer-reviewed program of grants and contracts. The Institute also conducts research in intramural NIH laboratories and clinical facilities and collaborates with other NIH institutes and Federal agencies.

Visit [www.nidcd.nih.gov](http://www.nidcd.nih.gov) for more information.



National  
Institute of  
Dental and  
Craniofacial  
Research

## MOLECULAR ASPECTS OF THE PERIPHERAL SENSORY SYSTEM

An estimated one-third of Americans suffer from chronic pain conditions, including cancer and musculoskeletal pain, trigeminal neuralgia, temporomandibular disorders, diabetic neuropathy, complex regional pain syndromes, and low back pain. The NIDCR supports basic, clinical, translational, and community-based research on the epidemiology, etiology, diagnosis, treatment, and prevention of inflammatory, neurodegenerative, and neuropathic conditions that lead to chronic pain or impaired oral function. The Institute also supports research on somatosensory and chemosensory function. Neurobiological studies on the trigeminal nerve system also fall within NIDCR's scope, as does research on the molecular genetic processes involved in injury, repair and regeneration within the trigeminal system. Through grant support and its intramural research program,

NIDCR currently sponsors research ranging from studies of neural development and regeneration to studies on the clinical and molecular aspects of pain perception.

NIDCR grantees and intramural researchers are studying the structural and molecular features of neural pathways that regulate pain sensation. Recent work by NIDCR-supported investigators demonstrated that one class of dental sensory neurons fails to develop in response to postnatal deprivation of nerve growth factor (NGF), a neurotrophic growth factor. Researchers are currently seeking to define the phenotype of distinct populations of sensory neurons that innervate dental pulp. Variation in the response to painful stimuli is the subject of translational research in NIDCR's Pain Research Clinic. Recent work has characterized wide variation in human phenotypes that may be related to single nucleotide polymorphisms (SNPs) in opiate receptors.

Clinical studies suggest that lesions of the trigeminal nerve lead to different chronic pain conditions from those caused by peripheral nerves because of the anatomical-physiological organization of the trigeminal system. NIDCR grantees are also examining the anatomical distribution and features of tooth innervation, including mechanisms involved in the response to different forms of nerve injury. Recent intramural work has identified a genetic locus in an animal model of neuropathic pain which may provide a target for novel therapeutic interventions for pain following nerve injury. Such findings illustrate the importance of understanding molecular-genetic mechanisms of nerve injury.

Ongoing work by NIDCR-funded investigators is shedding light on the molecular biology of pain transmission. An experimental mouse model with deletions of the genes encoding the neuropeptides substance P and neurokinin A was found to have reduced pain responsiveness to intensely noxious stimulation compared to control animals. Related studies show that the vanilloid receptor, an important link in the process of transducing noxious stimuli in the periphery into neuronal pain signals, has increased responsiveness under inflammatory conditions. This may explain, in part, why chronic inflammatory pain is difficult to treat.

Clinical studies being carried out in collaboration with the NIH's Pain and Palliative Care Program are evaluating novel strategies for selectively blocking neurons that signal chronic pain. Future challenges will include deciphering the molecular signals involved in nervous system development, repair, and regeneration.

Visit [www.nidcr.nih.gov](http://www.nidcr.nih.gov) for more information.



National  
Institute of  
Neurological  
Disorders  
and Stroke

## MOLECULAR SENSING AND PAIN

NINDS' interest in the molecular study of sensation covers the basic neurobiology of mechanoreceptors, proprioception, brainstem chemoreceptors, baroreceptors, but has primarily focused on pain. NINDS is the leading supporter of extramural pain research. Our broad portfolio includes research about the cellular and molecular mechanisms of peripheral, spinal, brainstem, and cortical pathways of pain. Recent advances in the cellular and molecular biology of ion channels, receptors, signal transduction cascades, and genetics of nociceptors, have led to significant new insights into the neurobiology of pain and sensation. The neuronal receptor for capsaicin, the chemical in chili peppers that causes pain, has recently been cloned, opening a new avenue for therapeutic exploration. Ongoing investigations of molecular approaches for pain treatments include targeting calcium channel subtypes, the tetrodotoxin insensitive sodium channel, glutamate receptors, substance P receptors, and cyclooxygenase.

nase. NINDS continues its long-term support of basic research upon which these approaches depend.

Visit [www.ninds.nih.gov](http://www.ninds.nih.gov) for more information.



National Eye  
Institute

## THE MOLECULAR BASIS OF VISUAL TRANSDUCTION

The NEI, as part of its mission to protect and prolong the vision of the American people, supports basic vision research.

The retina is a biological image processor. It transforms each visual image into electrical signals, codes these signals, and finally sends this code as patterns of electrical information to the brain.

Normal vision requires the retina to operate over illumination conditions that span ten orders of magnitude, a feat well beyond the capabilities of the most sophisticated optical instruments. Most of the known neurotransmitters, neuropeptides, and neurohormones are represented in the retina.

Biochemical, molecular biological, physiological, and structure-function studies have produced a detailed understanding of the phototransduction cascade. NEI-supported scientists have provided a molecular description of the key components of the pathway. This has now become a classic model that has led the way toward our current understanding of signal processing by neurotransmitter receptors and molecular sensors in other sensory systems and the brain.

Excited by light, rhodopsin activates transducin, a member of the GTP-binding protein superfamily which is found in the disc membrane of retinal rods. Activated transducin stimulates the activity of its effector, cGMP phosphodiesterase, leading to a decrease in intracellular levels of cGMP. Members of the G protein-coupled receptor family are central to signal processing. Different receptors in the family can discriminate between neurotransmitters, hormones, chemical odorants and even different wavelengths of light.

The recent mapping of the crystal structure of rhodopsin has provided insight into the photoactivation process that relays changes in the rhodopsin molecule to transducin. A high-resolution, three-dimensional structure of transducin obtained by X-ray crystallography has also been used to understand the molecular mechanism of phototransduction. Guanosine diphosphate bound to transducin is exchanged for guanosine triphosphate following photoactivation of rhodopsin. These studies have also provided a glimpse of how transducin subsequently interacts with phosphodiesterase to catalyze the hydrolysis of cGMP.

The end target of the visual cascade pathway, the cGMP-gated channel, has also been characterized at the molecular and physiological levels. These studies have provided important new insight into how the activity of this channel is controlled by cGMP and extracellular divalent cations and is modulated by calmodulin.

Advances in our understanding of vision and particularly visual transduction have yielded important new insights into the causes of retinal diseases. The majority of these diseases affect photoreceptors and its supporting pigment epithelium. For some inherited retinal diseases the affected gene and protein have been identified. NEI-supported scientists are beginning to understand the early effects on photoreceptors of abnormal proteins that have lost their function. This may lead to development of new therapies for these disorders and would not have been possible without the basic biological studies of vision and visual transduction.

Visit [www.nei.nih.gov](http://www.nei.nih.gov) for more information.



# Order, order

Some of the latest offerings for the burgeoning sequencing industry.

## KingFisher genomic DNA kit

**Thermo Hybaid** [www.thermohyaid.com](http://www.thermohyaid.com)

### Rapid DNA purification

This new kit is designed to purify high molecular weight genomic DNA from a variety of starting materials, including blood, tissue and cells, and can be used with the company's KingFisher or KingFisher mL magnetic particle processors. It utilizes silica particles for DNA capture in optimized reaction conditions. With the KingFisher processor, 24 samples can be purified in half-an-hour, or 15 samples with the KingFisher mL. Applications include PCR, creating databases of DNA samples and forensic analysis. The kit is available in sample sizes of 96 and 300 for the KingFisher and 30 and 120 for the KingFisher mL.

## Sequencher 4.0

**Gene Codes** [www.genecodes.com](http://www.genecodes.com)

### DNA sequencing software for Mac users

Sequencher software allows scientists to compare thousands of DNA sequence fragments with each other and to align them, arranging them in their proper orientation and order in contigs. Sequencher 4.0 for Macintosh features support for base-call confidence values, integration with Phrap files and automatic contig editing to match high-quality overlaps. Other new capabilities include comparison to a defined reference sequence, compact reporting of variations and SNPs, and additional support for forensic sequencing applications.

## GeneJumper

**Invitrogen** [www.invitrogen.com](http://www.invitrogen.com)

### Transposon primer insertion kit

Instead of subcloning DNA fragments into vectors with sequencing primer binding sites, GeneJumper inserts primer sites randomly into target DNA. Transformation of transposed plasmids generates thousands of sequencing templates for overlapping coverage with no need to synthesize custom primers or shotgun clones.

## SequiTherm EXCEL II

**Epicentre** [www.epicentre.com](http://www.epicentre.com)

### Sequencing kits that stay on track

EXCEL II DNA sequencing kits are suitable for sequencing a broad range of templates and can be used for general sequencing and finishing projects using primer and internal dye-labelling technologies. The kits can read through very difficult templates with regions of high GC content, inverted or direct repeats, dinucleotide repeats, homopolymeric stretch-

## MagAttract

**Qiagen**

[www.qiagen.com](http://www.qiagen.com)

### A miniprep system for the BioRobot 8000

MagAttract can be used on the BioRobot 8000 to purify plasmid DNA for high-throughput projects such as genome sequencing. MagAttract combines silica-based DNA purification with automated magnetic bead separation to allow rapid lysis of cells and direct purification of plasmid DNA without lysate-clearing steps. Ready-to-use protocols can process up to 384 samples (in four 96-well microplates) in parallel. The BioRobot 8000 high-speed shaker system resuspends magnetic beads in a variety of vessels, including 96-well plates and square-well blocks. Reproducible yields of up to 3 µg of plasmid DNA from 1.25 ml of bacterial culture can be achieved.



Feel the force: MagAttract for DNA purification.

es, hairpin loops and localized areas of inter-strand reannealing. Stops, ambiguous base calls and nonspecific background are reduced. EXCEL II kits are available for use with automated DNA sequencers or manual sequencing.

## ULTRArray-Advantage

**Ambion**

[www.ambion.com](http://www.ambion.com)

### Bound to be repeated

This system has two human ULTRArray membranes, each containing 8,400 unique human gene products. It also includes everything necessary to synthesize cDNA probes and to hybridize and wash the arrays. Ambion's Strip-EZ technology can be used to synthesize a StripAble probe. The mild stripping conditions used to remove the probe preserve the nucleic acid bound to the membrane, resulting in an array that can be used 10 or more times with no loss of signal.

## CEQ 2000XL

**Beckman Coulter** [www.beckmancoulter.com](http://www.beckmancoulter.com)

### Take a load off DNA sequencing

This system for fully automated DNA sequencing and fragment analysis uses capillary electrophoresis technology with a linear polyacrylamide gel to increase read lengths while shortening run times. Cycle time per capillary has been reduced to 100 minutes, reading out up to 700 bases. Plate washing, gel casting and tasks associated with loading samples onto slab gels are eliminated. The system features laser-induced fluorescence detection and four-colour dideoxym Terminator chemistry, and operates in a microplate format. In addition to

base calling, the software measures the accuracy of each base call, eliminating the need for a separate program for accuracy determination.

## TOP

**Varian**

[www.varianinc.com](http://www.varianinc.com)

### Top-notch purification

The TOP (trityl-on oligonucleotide purification) high-throughput system combines a 96-well plate with a streamlined protocol and a resin that will deliver optimum yield and purity for oligonucleotides up to 200 nmol in scale and 50-mers in length. It can purify up to 96 samples in 45 minutes and is fully compatible with liquid handling/robotic systems. The modular design offers a flexible 96-well plate format that can purify up to 96 samples simultaneously. No vacuum manifolds are needed, and the eight-step procedure uses a small solvent volume and requires only a single pass of the sample.

## RoboSmart

**MWG Biotech**

[www.mwg-biotech.com](http://www.mwg-biotech.com)

### Independent worker

RoboSmart can carry out automated, single-step plasmid or PCR purification, set-up, processing and subsequent purification of the sequencing reaction. The unit combines several process steps on a single platform. Once the robot is loaded, a task that can be completed in half-an-hour, more than 768 samples will be processed within 15 hours, with no further user intervention. RoboSmart can be used with most capillary sequencers to provide fully automated sample preparation.



## PRISM 7900HT

**Applied Biosystems** [www.appliedbiosystems.com](http://www.appliedbiosystems.com)

### High-throughput sequence detection

This real-time PCR instrument performs gene expression quantification and detection of SNPs at high throughputs by taking a higher plate density than previous instruments without compromising speed, resolution or reproducibility. The sensitive probe-based, fluorogenic 5' nuclease assay allows multiplex reactions for real-time quantification and SNP genotyping, and the SYBR Green 1 double-stranded DNA binding dye assay is suitable for target identification in screening assays. Laser light is distributed to each well in the reaction plate, exciting the fluorescent dyes present. The fluorescent light generated is then directed through a spectrograph to a CCD camera to yield spectrally resolved fluorescence data.

## Lasergene v5.0

**DNASTAR** [www.dnastar.com](http://www.dnastar.com)

### Software for Windows and Macs

This upgraded version of Lasergene sequence analysis software includes the ClustalW alignment algorithm in the extensive list of pairwise and multiple sequence alignment options offered by this program. Both fast-approximate and slow-accurate options are provided for ClustalW and, as with the other alignment methods, the full range of ClustalW parameters may be modified by the user. In addition, any subsection of an initial alignment can be selected and saved as a new project in order to refine alignments and improve phylogenetic analysis. Alignments may be saved in graphical format or as plain text. The Lasergene v5.0 sequence assembler now imports Phrap assemblies, simplifying the transition from Unix to desktop computers.

## SNPTwist and Prophecy ToolKit

**DoubleTwist** [www.doubletwist.com](http://www.doubletwist.com)

### A new twist on software for SNP detection

The SNPTwist program allows rapid, high-throughput detection of SNPs within large sequence databases. It incorporates DoubleTwist's SNP detection algorithm which enables mining of large and redundant data sets for SNPs. The program accepts data from multiple sources, allowing SNP detection to be combined with existing sequence analysis research programs. The Prophecy Toolkit is a data integration program that provides an optional enhancement for the DoubleTwist Annotated Human Genome Database and associated software for data interpretation and visualization. It consists of a suite of UNIX-based software programs that will integrate proprietary sequence data into the DoubleTwist database. Results can be merged, mined and visualized with a single command. Data types that can be merged include protein and nucleotide sequences and raw BLAST results.



The real world: sequence detection in real time, from the PRISM 7900HT.

## Rapid-Screen cDNA panel

**OriGene** [www.origene.com](http://www.origene.com)

### Functional genomics

OriGene offers customized screening of either a 12-tissue human or a six-tissue mouse Rapid-Screen Arrayed cDNA Library Panel (Super-Plate), and subsequent isolation of the 'longest clone' from the selected library. Since most genes are expressed in several tissues, the screening of a multi-tissue panel, containing a total of either 6-million human or 3-million mouse size-selected clones, maximizes the probability of isolating a full-length clone. The 12 human libraries currently available are brain (adult), brain (foetal), heart, kidney, liver, lung, muscle, peripheral blood leukocytes, placenta, small intestine, spleen and testis. The mouse libraries are brain (adult), embryo (12.5-day), embryo (19-day), liver, testis, and thymus.

## Fastlink

**Cambio** [www.cambio.co.uk](http://www.cambio.co.uk)

### Speedy ligation

The Cambio Fastlink DNA ligation and screening kit allows DNA ligation, transformation and screening of clones in less than 24 hours. Any insert can be ligated into plasmid, cosmid or lambda vectors. Ligation of sticky ends can be performed in five minutes at room temperature, with blunt ends ligated in 15 minutes, while PCR products with an A-overhang can be ligated in one hour at 16°C. High efficiencies are routinely achieved, and the ligation product does not need to be desalted prior to transformation of electro-competent or chemically competent cells.

*These notes are compiled in the Nature office from information provided by the manufacturers.*

## ADVERTISEMENTS

## CUSTOMISED immunoassays for research and production



**BioGenes**

Gesellschaft für Biopolymere mbH

Internet [www.biogenes.de](http://www.biogenes.de)

Berlin • Germany

Phone +49 (0)30-65 76 23 96

Look what's new...

## EDG (GPCR)

Human Recombinant GPCRs  
Membrane Preps  
-Lysophospholipid  
growth factor  
receptors

...From Exalpha Biologicals, Inc.

Find out more...

Phone: 617-445-6463 x 504

Website: [www.exalpha.com](http://www.exalpha.com)

**Exalpha**

| Proposed Nomenclature | New | Old   |
|-----------------------|-----|-------|
| LPA <sub>1</sub>      | =   | EDG 2 |
| LPA <sub>2</sub>      | =   | EDG 4 |
| LPA <sub>3</sub>      | =   | EDG 7 |
| SIP <sub>1</sub>      | =   | EDG 1 |
| SIP <sub>2</sub>      | =   | EDG 5 |
| SIP <sub>3</sub>      | =   | EDG 3 |
| SIP <sub>4</sub>      | =   | EDG 6 |
| SIP <sub>5</sub>      | =   | EDG 8 |

## DIVERSE SMALL MOLECULES READY FOR SCREENING



10K-50K  
Compounds  
in Plates

• Next Day Delivery\*

• Competitively Priced

\*(Some Restrictions Apply)

Chembridge Corporation • 800.964.6143

858.451.7400 • 858.451.7401 FAX • E-mail: [chem@chembridge.com](mailto:chem@chembridge.com)

[www.chembridge.com](http://www.chembridge.com)

## BACs • YACs • PACs cDNAs

We provide clones and colony membranes for a number of human, mouse and rat libraries as well as a custom screening service for the libraries.

**ResGen**

Invitrogen Corporation

U.S. or Canada 800-533-4363

FAX 256-536-9016

Check out our web site at  
[www.resgen.com](http://www.resgen.com)

## Contacts

**Publisher:** Fabien Savenay  
**Editor:** Paul Smaglik  
**Sales Director:** Ben Crowe

### European Head Office, London

The Macmillan Building  
4 Crinan Street  
London N1 9XW, UK  
Tel +44 (0) 20 7843 4961  
Fax +44 (0) 20 7843 4996  
e-mail: [classified@nature.com](mailto:classified@nature.com)

### Group European Manager:

Leonie Wells (4954)

### European Manager:

Nevin Bayoumi (4978)

### UK/ RoW/ Ireland:

Matt Powell (4953), Ben Corp (4974),  
Andy Douglas (4975)

**Holland/ Italy:** Nevin Bayoumi (4978)

**Scandinavia:** Silje Opstrup (4994)

**Spain/ Portugal:** Leonie Wells (4954)

### Production Manager:

Billie Franklin  
To send films and materials use  
London address above.  
Tel +44 (0) 20 7843 4814  
Fax +44 (0) 20 7843 4996  
e-mail: [classified@nature.com](mailto:classified@nature.com)

### International

#### Advertising Coordinator:

Laura Pearson (4977)

#### Naturejobs web development:

Tom Hancock

#### Naturejobs online production:

Ben Lund

### European Satellite Offices

#### France/ Belgium:

Christine Niox-Chateau  
Tel + 33 (0) 1 43 20 16 51  
Fax + 33 (0) 1 43 20 51 52  
e-mail: [c.nioxchateau@nature.com](mailto:c.nioxchateau@nature.com)

#### Germany/ Austria/ Switzerland:

Patrick Phelan/ Kate Turner  
Tel + 49 89 54 90 57 11/-2  
Fax + 49 89 54 90 57 20  
e-mails: [p.phelan@nature.com](mailto:p.phelan@nature.com)  
[k.turner@nature.com](mailto:k.turner@nature.com)

### US Head Office, New York

345 Park Avenue South,  
10th Floor, New York, NY 10010-1707  
Tel +1 800 989 7718  
Fax +1 800 989 7103  
e-mail: [classified@natureny.com](mailto:classified@natureny.com)

### US Sales Director:

Ben Crowe

### US Sales Manager:

Peyton Mason

### US Advertising Coordinator:

Corissa Salzman

### Japan Head Office, Tokyo

MG Ichigaya Building (5F),  
19-1 Harakatomachi,  
Shinjuku-ku,  
Tokyo 162-0841  
Tel +81 3 3267 8751  
Fax +81 3 3267 8746  
e-mail: [k.cowan@naturejpn.com](mailto:k.cowan@naturejpn.com)  
**Japan Manager:** Kate Cowan

# naturejobs

nature.com

## A rich diversity

**W**arren Bush always tells young scientists that their future career paths will bear little resemblance to the route that his own took. Bush, a volunteer career consultant for the American Chemical Society (ACS), worked for 40 years at just one large company — young chemists of today will probably work for several smaller firms during their working lives, he says. Bush's employer, Shell, focused on inorganic chemistry, but organic chemists are now much more in demand, Bush says. And he notes that Shell emphasized manufacturing — whereas, today, one of the biggest growth areas for chemists is likely to be in services.

But the changing jobs market-place doesn't mean that prospects are bleak, Bush told a group of job-seeking chemists at a careers workshop during the annual meeting of the ACS in Chicago last month. Quite the contrary. Unemployment among US chemists is only about 1.5%, and average salaries have risen steadily over the past 10 years at a rate well above inflation. "The job market is as good as it gets," Bush said.

But to take advantage of this positive climate, chemists will need to shop for jobs differently. Rather than hope to land a lifetime position at a large company, they may need to target many smaller firms, ideally landing several offers so that they can choose the one that most suits them at the time.

Bush noted that weighing these odds can be difficult. A company boasting a future 20% boost in recruitment may actually offer only one or two jobs to chemists because only a fraction of its workforce is actively engaged in science. And that fraction is increasingly divided among different disciplines and specialities.

This new market-place may mean more diverse opportunities. But it may also require more flexibility.

### Paul Smaglik

Naturejobs editor



## Contents

### MOVERS

Austin gains from Bell Labs;  
Renovis reworks its board;  
bioterrorism faces a fresh foe;  
and more [Back page](#)

### WWW.NATURE.COM/ NATUREJOBS

Career centre  
Information on the scientific  
job market

### FOCUS

### SPOTLIGHT

### RECRUITMENT

### SCIENTIFIC ANNOUNCEMENTS

### SCIENTIFIC EVENTS

## PHYSICS

**Ananth Dodabalapur** has left Bell Labs after 11 years to take up an endowed chair at the University of Texas at Austin. Dodabalapur says his decision to leave Bell was not related to the financial woes of its parent company Lucent. "I've been wanting to go to academia for a while," he says. He adds that Bell Labs has a long history of producing scientists who later leave for academia. His new position will allow the electrical and computer engineer to build up an interdisciplinary programme in organic electronics and photonics, which will also draw from chemistry and microelectronics — areas that he sees as strengths at the university.

After positions in Switzerland, Germany, Berkeley and Boston, physicist **Shuheng Pan** has moved to the University of Houston, Texas. Although he enjoyed aspects of all his previous appointments, Pan says that Houston offered him a position that was too good to pass up — a tenured position with an endowed chair and plenty of guaranteed support to pay for postdocs and infrastructure. "They are providing me with very good start-up funds, and very good operational support," Pan says. "I don't have to struggle for funding." Pan's work centres on the scanning tunnelling microscope (STM), and with the new funds, he will build an STM that can handle supercooled samples — a project that he relishes. "Every time I go to a new place I try to develop a new design, try to push it further," he says. He is especially keen to increase the STM's capacity to handle supercooled samples as this increases resolution and allows physicists to probe phenomena that only occur at extreme temperatures. Pan will bring one postdoc with him from Boston University and is also working to get one from his native China.

## BIOTECHNOLOGY

Genomics-based neuroscience start-up Renovis last month appointed co-founder **Corey Goodman** as chief executive and member of its board of directors. Goodman, a neuroscientist at the University of California, Berkeley, and an investigator with the Howard Hughes Medical Institute, previously served the company as an

adviser. Renovis' acting president and chief operating officer, **Lynne Zydowsky**, who also co-founded the company, will continue to serve as chief operating officer, but on 1 September also assumed the role of senior vice-president, business development. A third co-founder, **Tito Serafini**, previously a neurobiologist at the University of California, Berkeley, was named chief discovery scientist and vice-president for genomic research; and **Andrew Shyjan** was named vice-president, molecular biology. The San Francisco-based company uses gene-expression profiles in neuronal cell types to identify drug targets associated with neuronal regeneration, pain, behavioural disorders and psychiatric diseases.

## GOVERNMENT

**Scott Lillibridge** will head the bioterrorism initiative at the US Department of Health and Human Services. Lillibridge has been with the US Centers for Disease Control and Prevention (CDC) since 1990, and has led its Bioterrorism Preparedness and Response Program since 1998. In 1995, he led the US medical delegation to Japan after the gas attack that killed 10 people in the Tokyo subway. He also participated in the federal public-health assessment following the Oklahoma City bombing in 1995. He has worked in 14 nations on epidemiology and other public-health issues. Before joining the CDC, Lillibridge served in the Indian Health Service in Oklahoma and Arizona.

The National Center for Research Resources, part of the National Institutes of Health, last month named **Anthony Hayward** as associate director of clinical research. Hayward is currently professor of paediatrics, microbiology and immunology at the University of Colorado Health Sciences Center and associate director of the university's Pediatric General Clinical Research Center. In his new position, Hayward will oversee several clinical research initiatives, including efforts in pancreatic-islet cells and gene therapy. Hayward's scientific accomplishments include the first description of leukocyte-adhesion-molecule deficiency as a blood disease, and identification of a new form of severe combined immunodeficiency in the Navajo Indian population.



Shuheng Pan



Scott Lillibridge



Anthony Hayward

## CONTACTING US AT MOVERS

Please send any information on new appointments or job moves to: [naturejobseditor@naturedc.com](mailto:naturejobseditor@naturedc.com).